

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 02:27:51 ; Search time 197 seconds

(without alignments)
1455.202 Million cell updates/sec

Title: US-10-661-809a-13

Perfect score: 3225
Sequence: 1 MKQLKQWVTSTLLILPL.....GAVLLIAGVFPARRKENA 627

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq.8:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*
10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3225	100.0	627	2	AAV00118 Enterococ
2	3225	100.0	627	5	ABP43337 E faecali
3	3225	100.0	627	6	ABU08365 E. faecal
4	3225	100.0	627	6	ABU13616 Enterococ
5	3225	100.0	627	9	ADY38952 Novel Ent
6	3225	100.0	659	7	ADH88239 Enterococ
7	3202	99.3	625	8	AD084889 E faecali
8	3202	99.3	625	8	AD084889 E faecali
9	2892	89.7	560	2	AAV00119 Enterococ
10	2892	89.7	560	5	ABP43338 E faecali
11	2892	89.7	560	6	ABU08366 E. faecal
12	2892	89.7	560	6	ABU13617 Enterococ
13	2892	89.7	560	6	ADY38954 Novel Ent
14	2446	75.8	625	8	AD084895 E faecium
15	2446	75.8	625	8	AD084895 E faecium
16	2446	75.8	625	8	AD084895 E faecium
17	438.5	13.6	476	8	AD084887 E faecali
18	438.5	13.6	476	8	AD084887 E faecali
19	414.5	12.9	473	8	AD084883 E faecium
20	395	12.2	470	8	AD084893 E faecium
21	355	11.0	508	7	AD084893 E faecium
22	356	11.0	581	5	ABP23737 Streptoco
23	352	10.9	554	5	ABP29802 Streptoco

24	352	10.9	554	8	ADL00034 Streptoco
25	352	10.9	554	8	AD010463 Group B S
26	352	10.9	554	8	ADU69559 S agalact
27	352	10.9	554	8	ADV89776 Streptoco
28	352	10.9	554	8	ADV81029 Streptoco
29	352	10.9	554	8	ADV81029 Streptoco
30	352	10.9	554	9	ADZ02882 S. agalac
31	350	10.9	517	9	ADZ02883 S. agalac
32	347	10.8	430	2	ADZ02884 S. agalac
33	347	10.8	430	2	ADZ02884 S. agalac
34	347	10.8	430	6	ABU83393 E. faecal
35	347	10.8	430	6	ABU83393 E. faecal
36	347	10.8	430	6	ABU13644 Enterococ
37	334	10.4	502	4	ADY39008 Novel Ent
38	334	10.4	502	4	ADY39008 Novel Ent
39	333	10.3	705	5	ADY72357 Virulent
40	329.5	10.2	674	8	ADY72357 Virulent
41	329.5	10.2	674	8	ADY72357 Virulent
42	333	10.0	525	5	ADP65654 Bifidobac
43	312	9.7	525	9	ADZ02884 S. agalac
44	311	9.6	384	2	ADY00147 Enterococ
45	311	9.6	384	5	ABP43366 E faecali

ALIGNMENTS

RESULT 1	AAV00118 standard; protein; 627 AA.
ID	AAV00118
XX	AAV00118;
AC	AAV00118;
XX	AAV00118;
DT	20-APR-1999 (first entry)
XX	20-APR-1999 (first entry)
DE	Enterococcus faecalis protein EF058.
XX	Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
KM	Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
XX	detection; attenuation; antigenic.
OS	Enterococcus faecalis.
XX	Enterococcus faecalis.
PN	W09850554-A2.
XX	W09850554-A2.
PD	12-NOV-1998.
XX	12-NOV-1998.
PF	04-MAY-1998; 98WO-US008959.
XX	04-MAY-1998; 98WO-US008959.
PR	06-MAY-1997; 97US-0044031P.
XX	06-MAY-1997; 97US-0044031P.
PR	16-MAY-1997; 97US-0046655P.
XX	16-MAY-1997; 97US-0046655P.
PA	14-NOV-1997; 97US-0066009P.
XX	14-NOV-1997; 97US-0066009P.
PI	(HUMA-) HUMAN GENOME SCI INC.
XX	(HUMA-) HUMAN GENOME SCI INC.
DR	Kunesh CA, Choi GH, Bailey C, Hromocky J A;
XX	Kunesh CA, Choi GH, Bailey C, Hromocky J A;
DR	WPI, 1999-070095/06.
XX	WPI, 1999-070095/06.
PT	N-PSDB; AAV20108.
XX	N-PSDB; AAV20108.
PT	New isolated Enterococcus faecalis polynucleotides - used to develop
XX	New isolated Enterococcus faecalis polynucleotides - used to develop
PT	products for the detection of Enterococcus and for use in vaccines for
XX	products for the detection of Enterococcus and for use in vaccines for
PT	prevention or attenuation of Enterococcus infection.
XX	prevention or attenuation of Enterococcus infection.
PS	Claim 9, Page 133; 301pp; English.
XX	Claim 9, Page 133; 301pp; English.
CC	The present sequence represents a protein isolated from Enterococcus
XX	The present sequence represents a protein isolated from Enterococcus
CC	faecalis. The present invention describes genes, proteins and antigenic
XX	faecalis. The present invention describes genes, proteins and antigenic
CC	polypeptides isolated from E. faecalis. The proteins can be used in
XX	polypeptides isolated from E. faecalis. The proteins can be used in
CC	vaccines for preventing or attenuating an infection caused by a member of
XX	vaccines for preventing or attenuating an infection caused by a member of
CC	the Enterococcus genus in an animal. They can also be used for detecting
XX	the Enterococcus genus in an animal. They can also be used for detecting
CC	Enterococcus and isolating in a sample. The nucleotide sequences can be used
XX	Enterococcus and isolating in a sample. The nucleotide sequences can be used
CC	for detecting Enterococcus nucleic acids. Products from the present
XX	for detecting Enterococcus nucleic acids. Products from the present
CC	invention can also be used for screening compounds to identify agonists
XX	invention can also be used for screening compounds to identify agonists

CC and antagonists of *E. faecalis* protein activity
 XX Sequence 627 AA;

Query Match 100.0%; Score 3225; DB 2; Length 627;
 Best Local Similarity 100.0%; Pred. No. 7.9e-198;
 Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQOLKKWYTVSTLLILPLFTSVLGTTTAPFAENGESQVLIHKKKMTDLPPLIQNSG 60
 DB 1 MQOLKKWYTVSTLLILPLFTSVLGTTTAPFAENGESQVLIHKKKMTDLPPLIQNSG 60
 QY 61 KEMSEFDKXQGLADYVFSIYNTNEFEYRORAGASVDAKQAVOSLTPEKPVAAQGTDDAN 120
 DB 61 KEMSEFDKXQGLADYVFSIYNTNEFEYRORAGASVDAKQAVOSLTPEKPVAAQGTDDAN 120
 QY 121 GNVTVQLPKKQNGKDAVYTIKESPKGCVAAATNMVVAFPVYEMIKQTDGSYKGTBEELAV 180
 DB 121 GNVTVQLPKKQNGKDAVYTIKESPKGCVAAATNMVVAFPVYEMIKQTDGSYKGTBEELAV 180
 QY 181 VHIYPRNVAVANDGSLHVKVGTAEENGNGAFVLSKSGSPGVYKIQGVNDGLYTWTT 240
 DB 181 VHIYPRNVAVANDGSLHVKVGTAEENGNGAFVLSKSGSPGVYKIQGVNDGLYTWTT 240
 QY 241 DKEQAKRPITGKSYEIGENDPTEAENGTEGELTVKQLEVGSYLIEEVKADNNMELIENQTK 300
 DB 241 DKEQAKRPITGKSYEIGENDPTEAENGTEGELTVKQLEVGSYLIEEVKADNNMELIENQTK 300
 QY 301 TPEFTIANNQTVPEKTVKNDTSKVDKTPSLDGKVAIGEKIKYQISVNIPLGIADKEGD 360
 DB 301 TPEFTIANNQTVPEKTVKNDTSKVDKTPSLDGKVAIGEKIKYQISVNIPLGIADKEGD 360
 QY 361 ANKYVKFNLDVGDALTFDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVAANPAY 420
 DB 361 ANKYVKFNLDVGDALTFDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVAANPAY 420
 QY 421 IPTLPFGTLKFVYFMHLINEKADPTKGFKNKANVNDGHTDDQTPPVEVVTGGKRFIKYD 480
 DB 421 IPTLPFGTLKFVYFMHLINEKADPTKGFKNKANVNDGHTDDQTPPVEVVTGGKRFIKYD 480
 QY 481 GNVTVQALAGASPVVRDQNSDTANLKI DETTKAATWYKTAEAATFTTTADGLVDING 540
 DB 481 GNVTVQALAGASPVVRDQNSDTANLKI DETTKAATWYKTAEAATFTTTADGLVDING 540
 QY 541 LKGYTYLEETVAPDDYVLLTNRIEFVNVESYGTENTVSPKVPNKGKGLPSTGKG 600
 DB 541 LKGYTYLEETVAPDDYVLLTNRIEFVNVESYGTENTVSPKVPNKGKGLPSTGKG 600
 QY 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627
 DB 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627

RESULT 2
 ABP43337
 ID ABP43337 standard; protein; 627 AA.
 XX ABP43337;
 AC
 XX 05-AUG-2002 (first entry)
 DT
 XX E faecalis Bf058 protein.
 DE
 XX Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.
 KM
 XX Enterococcus faecalis.
 OS
 XX US2002045737-A1.
 PN
 XX 18-APR-2002.
 PD
 XX 04-MAY-1998; 98US-00071035.
 PF
 XX

US 10-661-809a-13

PR 04-MAY-1998; 98US-00071035.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Choi GH, Bailey C, Hromockyj A, Kunsch CA, G4480413
 PI WPI; 2002-425450/45.
 DR N-PSDB; ABN98093.
 XX
 PT New genes and polypeptides from *Enterococcus faecalis*, useful as vaccines
 for preventing, treating or attenuating an infection caused by a member
 of the *Enterococcus* genus in an animal, particularly *E. faecalis*.
 PS
 XX Claim 9; Page 93; 255pp; English.
 CC The present invention provides the protein and coding sequences of a
 CC number of polypeptides from *Enterococcus faecalis*. The proteins can be
 CC used as vaccines for preventing or attenuating an infection caused by a
 CC member of the *Enterococcus* genus in an animal, particularly *E. faecalis*.
 CC The polynucleotide is also useful for preventing or treating *E. faecalis*
 CC infection. The present sequence is a protein of the invention

QY Sequence 627 AA;
 XX
 XX Query Match 100.0%; Score 3225; DB 5; Length 627;
 Best Local Similarity 100.0%; Pred. No. 7.9e-198;
 Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQOLKKWYTVSTLLILPLFTSVLGTTTAPFAENGESQVLIHKKKMTDLPPLIQNSG 60
 DB 1 MQOLKKWYTVSTLLILPLFTSVLGTTTAPFAENGESQVLIHKKKMTDLPPLIQNSG 60
 QY 61 KEMSEFDKXQGLADYVFSIYNTNEFEYRORAGASVDAKQAVOSLTPEKPVAAQGTDDAN 120
 DB 61 KEMSEFDKXQGLADYVFSIYNTNEFEYRORAGASVDAKQAVOSLTPEKPVAAQGTDDAN 120
 QY 121 GNVTVQLPKKQNGKDAVYTIKESPKGCVAAATNMVVAFPVYEMIKQTDGSYKGTBEELAV 180
 DB 121 GNVTVQLPKKQNGKDAVYTIKESPKGCVAAATNMVVAFPVYEMIKQTDGSYKGTBEELAV 180
 QY 181 VHIYPRNVAVANDGSLHVKVGTAEENGNGAFVLSKSGSPGVYKIQGVNDGLYTWTT 240
 DB 181 VHIYPRNVAVANDGSLHVKVGTAEENGNGAFVLSKSGSPGVYKIQGVNDGLYTWTT 240
 QY 241 DKEQAKRPITGKSYEIGENDPTEAENGTEGELTVKQLEVGSYLIEEVKADNNMELIENQTK 300
 DB 241 DKEQAKRPITGKSYEIGENDPTEAENGTEGELTVKQLEVGSYLIEEVKADNNMELIENQTK 300
 QY 301 TPEFTIANNQTVPEKTVKNDTSKVDKTPSLDGKVAIGEKIKYQISVNIPLGIADKEGD 360
 DB 301 TPEFTIANNQTVPEKTVKNDTSKVDKTPSLDGKVAIGEKIKYQISVNIPLGIADKEGD 360
 QY 361 ANKYVKFNLDVGDALTFDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVAANPAY 420
 DB 361 ANKYVKFNLDVGDALTFDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVAANPAY 420
 QY 421 IPTLPFGTLKFVYFMHLINEKADPTKGFKNKANVNDGHTDDQTPPVEVVTGGKRFIKYD 480
 DB 421 IPTLPFGTLKFVYFMHLINEKADPTKGFKNKANVNDGHTDDQTPPVEVVTGGKRFIKYD 480
 QY 481 GNVTVQALAGASPVVRDQNSDTANLKI DETTKAATWYKTAEAATFTTTADGLVDING 540
 DB 481 GNVTVQALAGASPVVRDQNSDTANLKI DETTKAATWYKTAEAATFTTTADGLVDING 540
 QY 541 LKGYTYLEETVAPDDYVLLTNRIEFVNVESYGTENTVSPKVPNKGKGLPSTGKG 600
 DB 541 LKGYTYLEETVAPDDYVLLTNRIEFVNVESYGTENTVSPKVPNKGKGLPSTGKG 600
 QY 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627
 DB 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627

RESULT 3
ABU88365
ID ABU88365 standard; protein; 627 AA.
XX
XX ABU88365;
AC
XX
XX 07-JUL-2003 (first entry)
XX
XX
DE E. faecalis novel protein #109.
XX
XX Vaccine; endocarditis; bacteraemia; urinary tract infection; UTI;
XX intraabdominal infection; soft tissue infection; neonatal sepsis.
XX
XX Enterococcus faecalis.
OS
XX US2003017495-A1.
XX
XX 23-JAN-2003.
XX
XX 29-JUL-2002; 2002US-00206576.
XX
XX 06-MAY-1997; 97US-0044031P.
XX 16-MAY-1997; 97US-0046655P.
XX 14-NOV-1997; 97US-0066009P.
XX 04-MAY-1998; 98US-00071035.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Chai GH, Bailey C, Hromockyj A, Kunsch CA;
XX
XX WPI: 2003-416890/39.
XX N-PSDB; ACA88057.
XX
XX New nucleic acid molecules and polypeptides from Enterococcus faecalis,
XX useful as vaccines for preventing or attenuating an enterococcal
XX infection in an animal, or for identifying Enterococcus faecalis in
XX biological samples.
XX
XX Claim 12; Page: 40pp; English.
XX
XX The invention relates to a new isolated nucleic acid molecule comprising
XX a polynucleotide isolated from Enterococcus faecalis appearing as
XX ACA87949-ACA88196 (or sequences complementary to them or 95% identical to
XX them). Also included are the proteins encoded by the above nucleic acids,
XX making a recombinant vector (comprising inserting the isolated nucleic
XX acid molecule cited above into a vector), a host cell comprising the
XX vector (used to produce the protein), an isolated antibody specific for
XX the polypeptides, a hybridoma that produces the antibody, an isolated
XX polypeptide antigen comprising an amino acid sequence of an Enterococcus
XX faecalis epitope listed in the specification, a vaccine comprising one or
XX more E. faecalis polypeptides (and a pharmaceutical diluent, carrier or
XX excipient) where the polypeptide elicits protective antibodies in an
XX animal to a member of the genus Enterococcus; preventing or attenuating
XX an infection caused by a member of the genus Enterococcus in an animal
XX comprising administering to the animal the polypeptide and detecting
XX Enterococcus nucleic acids in a biological sample. The E. faecalis
XX nucleic acid molecules and polypeptides are useful as vaccines for
XX preventing or attenuating an enterococcal infection in an animal (e.g.
XX endocarditis, bacteraemia, urinary tract infection (UTI), intraabdominal
XX infection, soft tissue infection and neonatal sepsis). The polypeptides
XX are also useful for detecting Enterococcus aureus in immunosays, as
XX epitope tags, as molecular weight markers, or for generating antibodies
XX that specifically bind E. faecalis polypeptides. The nucleic acid
XX molecules are also useful as probes for gene mapping, or for identifying
XX E. faecalis in biological samples. The kit and methods are useful for
XX detecting Enterococcus antibodies or nucleic acid molecules in a
XX biological sample. The present sequence is a novel E. faecalis
XX polypeptide of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from the USPTO at
XX seqdata.uspto.gov/sequence.html?DocId=20030017495
XX
XX Sequence 627 AA;

US Pat
6,448,804

Query Match 100.0%; Score 3225; DB 6; Length 627;
Best Local Similarity 100.0%; Pred. No. 7.9e-198;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKQLKKWVYVSTLLILPLFTSVLGTTTAAEENGSAQVYHKKQMTDLPDPLIONSG 60
DB 1 MKQLKKWVYVSTLLILPLFTSVLGTTTAAEENGSAQVYHKKQMTDLPDPLIONSG 60
QY 61 KEMSEFDKXQGLADVTFSYNTNFEYBQRAAGASVDAKQVQSITPRKPAQGTDDN 120
DB 61 KEMSEFDKXQGLADVTFSYNTNFEYBQRAAGASVDAKQVQSITPRKPAQGTDDN 120
QY 121 GNVTVQLPKQNGKQAVVTIKKEPKBGVAAATNMVAFVPMYMIKQDGSYKXGTEBLAV 180
DB 121 GNVTVQLPKQNGKQAVVTIKKEPKBGVAAATNMVAFVPMYMIKQDGSYKXGTEBLAV 180
QY 181 VAIYPRNVVANDGSLHVKKVGTAEENGNGAEFVYSKESGSPGVKYIQGVKQGLYTWTT 240
DB 181 VAIYPRNVVANDGSLHVKKVGTAEENGNGAEFVYSKESGSPGVKYIQGVKQGLYTWTT 240
QY 241 DKEQAKRFTTGKSYEIGENDPTEANGTGLTVKULEVGSYILEEYKAPNNNELLENQTK 300
DB 241 DKEQAKRFTTGKSYEIGENDPTEANGTGLTVKULEVGSYILEEYKAPNNNELLENQTK 300
QY 301 TPFTTEANNQTPVEKTVKNDSKVDKTPSLDGKQVAGGEKIKYQISVNIPLGIADKEGD 360
DB 301 TPFTTEANNQTPVEKTVKNDSKVDKTPSLDGKQVAGGEKIKYQISVNIPLGIADKEGD 360
QY 361 ANKYVKFNLVDGHDALTPDNTSGEYAVAYLDGDTVAPENVQYTEQANGFTVAVNPAY 420
DB 361 ANKYVKFNLVDGHDALTPDNTSGEYAVAYLDGDTVAPENVQYTEQANGFTVAVNPAY 420
QY 421 IPTLPFGTGLKRVYFMHLEKADPTKGFENAVNDNGHTDDTPPTVEYVTSKGKPIKXD 480
DB 421 IPTLPFGTGLKRVYFMHLEKADPTKGFENAVNDNGHTDDTPPTVEYVTSKGKPIKXD 480
QY 481 GGVTAQALAGASFVVRDQNSDTANYLKIDETTKAATWTKAABATFTPTTADGLVDING 540
DB 481 GGVTAQALAGASFVVRDQNSDTANYLKIDETTKAATWTKAABATFTPTTADGLVDING 540
QY 541 LKGYTYLSEETVAPDDVYLLTNRIFEVNVBQSYGTENLVSEKVPNNKHGTLPTSGKG 600
DB 541 LKGYTYLSEETVAPDDVYLLTNRIFEVNVBQSYGTENLVSEKVPNNKHGTLPTSGKG 600
QY 601 IYVYLGSGAVLLILIGVTFARRRKENA 627
DB 601 IYVYLGSGAVLLILIGVTFARRRKENA 627
RESULT 4
ABU13616
ID ABU13616 standard; protein; 627 AA.
XX
XX ABU13616;
AC
XX
XX 26-FEB-2003 (first entry)
XX
XX Enterococcus faecalis EF040 polypeptide #109.
XX
XX EF040; immunostimulant; antibacterial; gene mapping.
XX
XX Enterococcus faecalis.
OS
XX US6448043-B1.
XX
XX 10-SEP-2002.
XX
XX 04-MAY-1998; 98US-00071035.
XX
XX 06-MAY-1997; 97US-0044031P.
XX 16-MAY-1997; 97US-0046655P.
XX 14-NOV-1997; 97US-0066009P.
XX

```
PR 14-NOV-1997; 97US-0066099P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Chai GH, Bailey C, Hromockyj A, Kunsch CA;
XX
XX MPI; 2003-089120/08.
XX
XX N-PSDB; ABX61663.
XX
XX New EF040 polypeptides and polynucleotides from Enterococcus faecalis,
XX useful for generating an immune response against E. faecalis and other
XX Enterococcus species, and as vaccines against other bacterial genera.
XX
XX Example 1; Col 115-116; 146pp; English.
XX
XX The invention relates to polynucleotide fragments of a gene from
XX Enterococcus faecalis, EF040, and the polypeptides encoded by them. The
XX polypeptides are useful in detecting E. faecalis, as epitope tags, as
XX molecular weight markers on SDS-PAGE gels or for molecular sieve gel
XX filtration columns, in generating antibodies that specifically bind to
XX the E. faecalis polypeptides, in generating an immune response against E.
XX faecalis and other Enterococcus species and as vaccines against other
XX bacterial genera. The polynucleotides are useful as probes for gene
XX mapping and for identifying E. faecalis in biological samples. Sequences
XX ABU13508-ABU13755 represent EF040 polypeptides of the invention. Note:
XX The sequence data for this patent can also be obtained from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 627 AA;
SQ
Query Match 100.0%; Score 3225; DB 6; Length 627;
Best Local Similarity 100.0%; Pred. No. 7.9e-198;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKQLKKWYVTSLLILPLFTSVIGTTTAPAEENGESAQVYIHKKKMTDLPDPLIQNSG 60
DB 1 MKQLKKWYVTSLLILPLFTSVIGTTTAPAEENGESAQVYIHKKKMTDLPDPLIQNSG 60
QY 61 KEMSEFDKXQGLADVTFSIYVNTNEFYEQRAAGASVDAKQAVQSLTPGKPAQGTDTAN 120
DB 61 KEMSEFDKXQGLADVTFSIYVNTNEFYEQRAAGASVDAKQAVQSLTPGKPAQGTDTAN 120
QY 121 GNVTVQLPRKQNGKDAVYTIKEEPKEGVAAATNMVVAFPVYEMIKQDGSYKYGTEELAV 180
DB 121 GNVTVQLPRKQNGKDAVYTIKEEPKEGVAAATNMVVAFPVYEMIKQDGSYKYGTEELAV 180
QY 181 VHIYKRVNVANDGSLHVKVGTAEENGGLNGAEFVYSKSESGPGYVXYIOGVXDGLYTWT 240
DB 181 VHIYKRVNVANDGSLHVKVGTAEENGGLNGAEFVYSKSESGPGYVXYIOGVXDGLYTWT 240
QY 241 DKEQAKRFTGSKSYEIGENDPTEAENGTELGTVKLEVGSYLLEVKAPNNMELLENQTK 300
DB 241 DKEQAKRFTGSKSYEIGENDPTEAENGTELGTVKLEVGSYLLEVKAPNNMELLENQTK 300
QY 301 TPFTEIANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKIRYQISVNIPLGIADKED 360
DB 301 TPFTEIANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKIRYQISVNIPLGIADKED 360
QY 361 ANKYKRVNVKNDALFTDNVTSGRYAYALYDGTVAIPENYQTEQANGTVAVNPNY 420
DB 361 ANKYKRVNVKNDALFTDNVTSGRYAYALYDGTVAIPENYQTEQANGTVAVNPNY 420
QY 421 IPTLPFGGLKFEVYFMHLNEKADPTKGFNGEAVNDGHTDDQTPPVEVVTGSKAFIKYD 480
DB 421 IPTLPFGGLKFEVYFMHLNEKADPTKGFNGEAVNDGHTDDQTPPVEVVTGSKAFIKYD 480
QY 481 GDNVTAQALAGASFYVRDQNSDPTANYLKIDETTKATVWTKRAEATTTTADGLVDING 540
DB 481 GDNVTAQALAGASFYVRDQNSDPTANYLKIDETTKATVWTKRAEATTTTADGLVDING 540
QY 541 LKYGTYVLEETVAPDDYVLLINRIEFVNVESGYGTENIVSPBKIPNKKGLPSTGKG 600
DB 541 LKYGTYVLEETVAPDDYVLLINRIEFVNVESGYGTENIVSPBKIPNKKGLPSTGKG 600
```

```
QY 601 IYVYIGSGAVLLIAGVYFARRRKENA 627
DB 601 IYVYIGSGAVLLIAGVYFARRRKENA 627
RESULT 5
ADY38952
ID ADY38952 standard; protein; 627 AA.
XX
XX ADY38952;
XX
XX 05-MAY-2005 (first entry)
XX
XX Novel Enterococcus faecalis protein sequence SeqID218.
XX
XX protein purification; DNA purification; antibacterial; vaccine;
XX enterococcus faecalis infection.
XX
XX Enterococcus faecalis.
XX
XX US2005043528-A1.
XX
XX 24-FEB-2005.
XX
XX 06-AUG-2004; 2004US-00912362.
XX
XX 06-MAY-1997; 97US-0044031P.
XX
XX 16-MAY-1997; 97US-004655P.
XX
XX 14-NOV-1997; 97US-0066099P.
XX
XX 04-MAY-1998; 98US-00071035.
XX
XX 29-JUL-2002; 2002US-00206576.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Chai GH, Bailey C, Hromockyj A, Kunsch CA;
XX
XX MPI; 2005-180840/19.
XX
XX N-PSDB; ADY38951.
XX
XX New isolated Enterococcus faecalis nucleic acid useful for diagnosing,
XX preventing and/or attenuating infection by Enterococcus faecalis in
XX animals, particularly in humans.
XX
XX Claim 9; SEQ ID NO 218; 28pp; English.
XX
XX This invention relates to novel genes and the proteins they encode
XX isolated from Enterococcus faecalis. The invention may be useful for the
XX development of compound with an antibacterial activity or a vaccine. The
XX present invention is useful for the diagnosis, prevention and/or
XX attenuation of infection by Enterococcus faecalis in animals,
XX particularly in humans. The present sequence is that of a novel
XX Enterococcus faecalis protein of the invention. Note: The present
XX sequence does not form part of the printed specification but was obtained
XX in electronic form from the USPTO web.
XX
XX Sequence 627 AA;
SQ
Query Match 100.0%; Score 3225; DB 9; Length 627;
Best Local Similarity 100.0%; Pred. No. 7.9e-198;
Matches 627; Conservative 100.0%; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKQLKKWYVTSLLILPLFTSVIGTTTAPAEENGESAQVYIHKKKMTDLPDPLIQNSG 60
DB 1 MKQLKKWYVTSLLILPLFTSVIGTTTAPAEENGESAQVYIHKKKMTDLPDPLIQNSG 60
QY 61 KEMSEFDKXQGLADVTFSIYVNTNEFYEQRAAGASVDAKQAVQSLTPGKPAQGTDTAN 120
DB 61 KEMSEFDKXQGLADVTFSIYVNTNEFYEQRAAGASVDAKQAVQSLTPGKPAQGTDTAN 120
QY 121 GNVTVQLPRKQNGKDAVYTIKEEPKEGVAAATNMVVAFPVYEMIKQDGSYKYGTEELAV 180
DB 121 GNVTVQLPRKQNGKDAVYTIKEEPKEGVAAATNMVVAFPVYEMIKQDGSYKYGTEELAV 180
```



```

OY 181 VHIYKRVVANDGSLHVKKVGTAEENGNGAEFVISKSESGPGTVKYIGVXKGLYTWTT 240
    |||||||
DB 181 VHIYKRVVANDGSLHVKKVGTAEENGNGAEFVISKSESGPGTVKYIGVXKGLYTWTT 240
OY 241 DKEQAKRPTGSKSYEIGENDPTEAENGTELVKNLEVGSYLIEVKAPNNMELIENQTK 300
    |||||||
DB 241 DKEQAKRPTGSKSYEIGENDPTEAENGTELVKNLEVGSYLIEVKAPNNMELIENQTK 300
OY 301 TPEFTIANNQTPVEKTVKNDTSKVDKTTTSLDCKDVAIGEKIKYQISVNIPLGIADKEGD 360
    |||||||
DB 301 TPEFTIANNQTPVEKTVKNDTSKVDKTTTSLDCKDVAIGEKIKYQISVNIPLGIADKEGD 360
OY 361 ANKYVKNLVNDKHDALTFEDNVTSGEYAVALYDGDVIAPENYQTEQANGFTVAVNPAY 420
    |||||||
DB 361 ANKYVKNLVNDKHDALTFEDNVTSGEYAVALYDGDVIAPENYQTEQANGFTVAVNPAY 420
OY 421 IPTLTPEGTLKRVYFMHLENEKADPTKGFKNENAVNDGHTDDQTPPVEVVTGGKRFIKVD 480
    |||||||
DB 421 IPTLTPEGTLKRVYFMHLENEKADPTKGFKNENAVNDGHTDDQTPPVEVVTGGKRFIKVD 480
OY 481 GGVTAIOALAGASFVVRDONSPTANYLKI DETTKAATWVKTA EATTTFTTADGLVDITG 540
    |||||||
DB 481 GGVTAIOALAGASFVVRDONSPTANYLKI DETTKAATWVKTA EATTTFTTADGLVDITG 540
OY 541 LKYGTYLLEETVAPDDVYLLTNRIEFVNVESGYTTEENVSPBEKVPNKHGTLPTSGGKG 600
    |||||||
DB 541 LKYGTYLLEETVAPDDVYLLTNRIEFVNVESGYTTEENVSPBEKVPNKHGTLPTSGGKG 600
OY 601 IYVYLGSGAVLLLLIAGVYFARRRKENA 627
    |||||||
DB 601 IYVYLGSGAVLLLLIAGVYFARRRKENA 627

```

RESULT 6

ADH88239 standard; protein; 659 AA.

ADH88239;

22-APR-2004 (first entry)

Enterococcus faecalis polypeptide #2719.

Enterococcus faecalis infection; transcription regulatory element;

antibacterial.

Enterococcus faecalis.

US6617156-B1.

09-SEP-2003

13-AUG-1998; 98US-00134000.

15-AUG-1997; 97US-0055778P.

(DOUC/) DOUCETTE-STAMM L A.

(BUSH/) BUSH D.

Doucette-Stamm LA, Bush D;

WPI; 2003-895394/82.

N-PBDB; ADH884834.

New nucleic acid comprising a sequence encoding an Enterococcus faecalis

polypeptide, useful for preparing a composition for diagnosing or

treating E. faecalis infection.

Disclosure; SEQ ID NO 6124; 193pp; English.

The invention relates to Enterococcus faecalis polynucleotides and

vector comprising a polynucleotide operably linked to a transcription
 CC regulatory element, a cell comprising a recombinant vector, a method for
 CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising
 CC a sequence not given in the specification, a recombinant vector
 CC comprising the nucleic acid and a cell comprising the recombinant vector.
 CC The polynucleotides can be used to detect the presence of E. faecalis in
 CC a sample. The sequences are useful for preparing a composition for
 CC diagnosing or treating Enterococcus faecalis infection. This sequence
 CC represents an E. faecalis polypeptide of the invention.

SQ Sequence 659 AA;

Query Match 100.0%; Score 3225; DB 7; Length 659;

Best local similarity 100.0%; Pred. No. 8,5e-198; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MKQLKKVWYVSTLLIILPFTSVIGTTPAFAENGESQVLVHKKMTDLPDPLIQNSG 60
    |||||||
DB 33 MKQLKKVWYVSTLLIILPFTSVIGTTPAFAENGESQVLVHKKMTDLPDPLIQNSG 92
    |||||||
OY 61 KEMSEPDKYOGLADVTPSIVYNTNEPFEORAGASVDAKQAVOSLTPGKPAQCTTDAN 120
    |||||||
DB 93 KEMSEPDKYOGLADVTPSIVYNTNEPFEORAGASVDAKQAVOSLTPGKPAQCTTDAN 152
    |||||||
OY 121 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAFVYEMIKQTDGSKYKTEBLAV 180
    |||||||
DB 153 GNVTVQLPKQNGKDAVYTIKEEPKEGVAAATNMVAFVYEMIKQTDGSKYKTEBLAV 212
    |||||||
OY 181 VHIYKRVVANDGSLHVKKVGTAEENGNGAEFVISKSESGPGTVKYIGVXKGLYTWTT 240
    |||||||
DB 213 VHIYKRVVANDGSLHVKKVGTAEENGNGAEFVISKSESGPGTVKYIGVXKGLYTWTT 272
    |||||||
OY 241 DKEQAKRPTGSKSYEIGENDPTEAENGTELVKNLEVGSYLIEVKAPNNMELIENQTK 300
    |||||||
DB 273 DKEQAKRPTGSKSYEIGENDPTEAENGTELVKNLEVGSYLIEVKAPNNMELIENQTK 332
    |||||||
OY 301 TPEFTIANNQTPVEKTVKNDTSKVDKTTTSLDCKDVAIGEKIKYQISVNIPLGIADKEGD 360
    |||||||
DB 333 TPEFTIANNQTPVEKTVKNDTSKVDKTTTSLDCKDVAIGEKIKYQISVNIPLGIADKEGD 392
    |||||||
OY 361 ANKYVKNLVNDKHDALTFEDNVTSGEYAVALYDGDVIAPENYQTEQANGFTVAVNPAY 420
    |||||||
DB 393 ANKYVKNLVNDKHDALTFEDNVTSGEYAVALYDGDVIAPENYQTEQANGFTVAVNPAY 452
    |||||||
OY 421 IPTLTPEGTLKRVYFMHLENEKADPTKGFKNENAVNDGHTDDQTPPVEVVTGGKRFIKVD 480
    |||||||
DB 453 IPTLTPEGTLKRVYFMHLENEKADPTKGFKNENAVNDGHTDDQTPPVEVVTGGKRFIKVD 512
    |||||||
OY 481 GGVTAIOALAGASFVVRDONSPTANYLKI DETTKAATWVKTA EATTTFTTADGLVDITG 540
    |||||||
DB 513 GGVTAIOALAGASFVVRDONSPTANYLKI DETTKAATWVKTA EATTTFTTADGLVDITG 572
    |||||||
OY 541 LKYGTYLLEETVAPDDVYLLTNRIEFVNVESGYTTEENVSPBEKVPNKHGTLPTSGGKG 600
    |||||||
DB 573 LKYGTYLLEETVAPDDVYLLTNRIEFVNVESGYTTEENVSPBEKVPNKHGTLPTSGGKG 632
    |||||||
OY 601 IYVYLGSGAVLLLLIAGVYFARRRKENA 627
    |||||||
DB 633 IYVYLGSGAVLLLLIAGVYFARRRKENA 659
    |||||||

```

RESULT 7

AD084889 standard; protein; 625 AA.

AD084889;

29-JUL-2004 (first entry)

E faecalis surface anchored LPXTG protein Seqid13.

LPXTG; cell wall-anchored surface protein; Gram positive bacterium;

extracellular matrix molecule; sequence database; C-terminal;

XX Enterococcus faecalis.

~~WO2004025456~~-A2.

25-MAR-2004.

15-SEP-2003; 2003WO-US028789.

13-SEP-2002; 2002US-0410303P.

(TEXA) UNIV TEXAS A & M SYSTEM.

(UABR-) UAB RES FOUND.

Hook M, Xu Y, Sillar

WDT 2004 21ECC04/200

N-PSDB; ADO84888.

Identifying LPTXs-containing cell wall-anchored surface proteins from Gram positive bacteria, for treating infection caused by the bacteria, comprises searching sequence information database for the sequence having LPTXG-motif.

Claim 16; SEQ ID NO 13; 96pp; English.

This invention relates to a novel method of identifying LpxTg-containing cell wall-anchored surface proteins from Gram positive bacteria that bind to an extracellular matrix molecule which comprises searching a database of sequence information for a putative protein sequence having the LpxTg-motif in its C-terminal region and analysing the sequence for the presence of one or more Immunoglobulin (Ig)-like fold regions. The invention may be useful for the production of compounds with an antibacterial activity or for production of a vaccine. In addition the disclosed sequences may be useful for gene therapy. The antibody is useful for treating or preventing an infection of Gram-positive bacteria in a human or animal patient. The method and the proteins are useful in generating antibodies for treating and preventing the spread of infections of Gram positive bacteria, for interfering with, or inhibiting binding interactions by Gram positive bacteria, for monitoring the level of gram positive bacterial antigens, or antibodies recognising the antigens in a human or animal patients suspected of containing the antigens or antibodies, in preventing or reducing infection of medical devices and prostheses caused by such organisms, and in treating or preventing infections in highly susceptible groups such as premature newborns, AIDS and debilitated cancer patients, and bone marrow transplantation. The present sequence is that of a surface anchored LpxTg protein identified using the method of the invention.

SQ Sequence 625 AA;

Query Match	99.3%	Score 3202;	DB 8;	Length 625;
Best Local Similarity	99.7%	Pred. No. 2.3e-196;		
Matches 625; Conservative	0;	Mismatches 0;	Indels 2;	Gaps 1

QY	1	MOLKQKWWTVSTLLILLEPLFTSVLGTTTAPAEENESQVLVHKKKMTDLPPLQNSG	60
Db	1	MOLKQKWWTVSTLLILLEPLFTSVLGTTTAPAEENESQVLVHKKKMTDLPPLQNSG	60
QY	61	KEMSEFDKQGLADVYFSIYNVTNFEYEOBAAQASVDAKAQVSLTPGKPAQGTDDAN	120
Db	61	KEMSEFDKQGLADVYFSIYNVTNFEYEOBAAQASVDAKAQVSLTPGKPAQGTDDAN	120
QY	121	GNVTVOLPKKQNGKDAVYTIKEEPKKGVAATMMVAFPYEMIKQTDGSKYKGTETLAV	180
Db	121	GNVTVOLPKKQNGKDAVYTIKEEPKKGVAATMMVAFPYEMIKQTDGSKYKGTETLAV	180
QY	181	VHIIYPRNVVANDGSLHVKKVGTVAENEGNGLNGAEFVISKSGSPGTVKTIQGVKQGLYTWTT	240

Dd	181	VH1PKNVAVNDGSL - KKVGTAENBGLNAGAFV1SKSGSDGTVCX1QGVNDGLYTWTF	238
Qy	241	DKEQAKRFITGKSVEIGENDPTEANGGELTVKQLEVGSY1LLEEKAPRNNAEL1ENQK	300
Dd	239	DKEQAKRFITGKSVEIGENDPTEANGGELTVKQLEVGSY1LEEKAPRNNAEL1ENQK	298
Qy	301	TPFTLEANNQRPVEKTVKNDTSKVDKTPSLDGKQVAGGEK1KYQ1SVN1PG1ADKEGD	366
Dd	299	TPFTLEANNQRPVEKTVKNDTSKVDKTPSLDGKQVAGGEK1KYQ1SVN1PG1ADKEGD	358
Qy	361	ANKYVKFNLVDGHDALTFTDNTSGEYAYALVDGPTV1APENYQVTEQANGFTVANPAY	420
Dd	359	ANKYVKFNLVDGHDALTFTDNTSGEYAYALVDGPTV1APENYQVTEQANGFTVANPAY	418
Qy	421	IPTLPGSTLKFVYFMH1NEKADPTKGFKNEANVNDGHTDDQTPPTVEVVTGGKRF1KYD	480
Dd	419	IPTLPGSTLKFVYFMH1NEKADPTKGFKNEANVNDGHTDDQTPPTVEVVTGGKRF1KYD	478
Qy	481	GDVTAQALAGASFVVRDONSPTANYLK1DEFTTKAATVYK1QAETFTPTTADGLVDITG	540
Dd	479	GDVTAQALAGASFVVRDONSPTANYLK1DEFTTKAATVYK1QAETFTPTTADGLVDITG	538
Qy	541	LKGYTYL1EETVAPDDYVLL1TR1R1FVNVNOSVGTENT1VSPKVPNNKKGL1PSTGGK	600
Dd	539	LKGYTYL1EETVAPDDYVLL1TR1R1FVNVNOSVGTENT1VSPKVPNNKKGL1PSTGGK	598
Qy	601	IYVYLSGAVLL1L1AGVYFARRKENA 627	
Dd	599	IYVYLSGAVLL1L1AGVYFARRKENA 625	

RESULT 8

ID ADV16503 standard; protein; 625 AA

AC ADV16503;

DT 24-FEB-2005 (first entry)

DE E. faecalis V583 hyperimmune serum reactive antigen protein - SEQ ID 198.
XX
XX antigen; antibacterial; vaccine; enterococcus infection; infection;
KM pharyngitis; impetigo; rheumatic fever; antipyretic; antirheumatic;
KM immunosuppressive; sepsis.

OS Enterococcus faecalis V583.

FT	Key	Location/Qualifiers
FT	Region	4. .28
FT		/note="Serum reactive epitope; claimed"
FT	Region	39. .45
FT		/note="Serum reactive epitope; claimed"
FT	Region	52. .58
FT		/note="Serum reactive epitope; claimed"
FT	Region	69. .82
FT		/note="Serum reactive epitope; claimed"
FT	Region	93. .115
FT		/note="Serum reactive epitope; claimed"
FT	Region	98. .127
FT		/note="Serum reactive epitope; claimed"
FT	Region	98. .123
FT		/note="Serum reactive epitope; claimed"
FT	Region	104. .136
FT		/note="Serum reactive epitope; claimed"
FT	Region	122. .138
FT		/note="Serum reactive epitope; claimed"
FT	Region	135. .140
FT		/note="Serum reactive epitope; claimed"
FT	Region	146. .163
FT		/note="Serum reactive epitope; claimed"
FT	Region	177. .192
FT		/note="Serum reactive epitope; claimed"

FT Region 209..215 /note= "Serum reactive epitope; claimed"
FT Region 221..232 /note= "Serum reactive epitope; claimed"
FT Region 271..284 /note= "Serum reactive epitope; claimed"
FT Region 331..337 /note= "Serum reactive epitope; claimed"
FT Region 341..352 /note= "Serum reactive epitope; claimed"
FT Region 360..378 /note= "Serum reactive epitope; claimed"
FT Region 383..390 /note= "Serum reactive epitope; claimed"
FT Region 392..401 /note= "Serum reactive epitope; claimed"
FT Region 409..422 /note= "Serum reactive epitope; claimed"
FT Region 428..435 /note= "Serum reactive epitope; claimed"
FT Region 462..470 /note= "Serum reactive epitope; claimed"
FT Region 474..480 /note= "Serum reactive epitope; claimed"
FT Region 482..496 /note= "Serum reactive epitope; claimed"
FT Region 531..539 /note= "Serum reactive epitope; claimed"
FT Region 541..549 /note= "Serum reactive epitope; claimed"
FT Region 551..560 /note= "Serum reactive epitope; claimed"
FT Region 562..569 /note= "Serum reactive epitope; claimed"
FT Region 576..582 /note= "Serum reactive epitope; claimed"
FT Region 598..618 /note= "Serum reactive epitope; claimed"
FT Region /note= "Serum reactive epitope; claimed"
FN WO2004106367-A2.
XX
XX PD 09-DEC-2004.
XX
XX PP 26-MAY-2004; 2004WO-EP005664.
XX
XX PR 30-MAY-2003; 2003EP-00450137.
XX
XX PA (INTE-) INTERCELL AG.
XX
XX PI Meinke A, Nagy E, Hanner M, Gelbmann D;
XX
XX DR WPI, 2005-039707/04.
XX
XX DR N-PSDB; ADV16333.
XX
XX PT Novel isolated nucleic acid molecule encoding hyperimmune serum reactive
XX PT antigen e.g., EP0020, EP0062, EP0149, EP0196, EP0253, EP0270,
XX PT EP0298 or its fragment, useful for producing vaccine against enterococcal
XX PT infection.
XX
XX PS Example 3; SEQ ID NO 198; 175pp; English.
XX
XX CC The invention relates to a novel isolated nucleic acid molecule encoding
XX CC a hyperimmune serum reactive antigen e.g. SEQ ID 171-340, 357-372 or 425-
XX CC 476, or its fragment. The antigen of the invention demonstrates
XX CC antibacterial activities and may be useful for producing a pharmaceutical
XX CC preparation, particularly a vaccine, against enterococcal infection.
XX CC Conditions associated with bacterial infection which may be prevented or
XX CC treated include bacterial pharyngitis, scarlet fever, impetigo, rheumatic
XX CC fever, necrotizing fasciitis and sepsis in humans. The current sequence
XX CC is that of an Enterococcus faecalis V583 hyperimmune serum reactive
XX CC antigen protein of the invention.
XX
XX SQ Sequence 625 AA;

Query Match 99.3%; Score 3202; DB 9; Length 625;
Best Local Similarity 99.7%; Pred. No. 2,3e-196;
Matches 625; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
QY 1 MKQLKKWYVSTLLILPLFTSVLGTTPAFAENGESAQVLIHKKMTDLPDPLIQNSG 60
DB 1 MKQLKKWYVSTLLILPLFTSVLGTTPAFAENGESAQVLIHKKMTDLPDPLIQNSG 60
QY 61 KEMSEPDKQGLADYTPSYNTNFEYBQRAAGASVDAKQAVQSLTPKPAQGTDDN 120
DB 61 KEMSEPDKQGLADYTPSYNTNFEYBQRAAGASVDAKQAVQSLTPKPAQGTDDN 120
QY 121 GNVTVQLPKKQKQAVNTIKKPKKGVAAATNMVAPFVPMIKQTDSSYKGTBELAV 180
DB 121 GNVTVQLPKKQKQAVNTIKKPKKGVAAATNMVAPFVPMIKQTDSSYKGTBELAV 180
QY 181 VHIYPRNVVANDSLHVKKVGTAENEGNGAFAVSKSBSGPGTVKYIQVGKGLYTWTT 240
DB 181 VHIYPRNVVANDSLHVKKVGTAENEGNGAFAVSKSBSGPGTVKYIQVGKGLYTWTT 240
QY 241 DKEQAKRPITGKSYTIGENDPTEANGGELTVKULEFGSYLLEEVKAPNNABELIENQK 300
DB 241 DKEQAKRPITGKSYTIGENDPTEANGGELTVKULEFGSYLLEEVKAPNNABELIENQK 300
QY 301 TPFTTBANNQTPVEKTVKNDTSKVDKTPSLDGKQVAGKIKYQISVNIPLGIADKED 360
DB 301 TPFTTBANNQTPVEKTVKNDTSKVDKTPSLDGKQVAGKIKYQISVNIPLGIADKED 360
QY 361 ANKYVKFNLVDGHDALTFDNTSGEYAVALYDGGTVAPENYQVTEQANGFTVAVNPAY 420
DB 361 ANKYVKFNLVDGHDALTFDNTSGEYAVALYDGGTVAPENYQVTEQANGFTVAVNPAY 420
QY 421 IPTLTGPGTLKRVYFMHLEKADPTKGRNEANVDNGHTDDQTPPVEVYVYTGKRPVKD 480
DB 421 IPTLTGPGTLKRVYFMHLEKADPTKGRNEANVDNGHTDDQTPPVEVYVYTGKRPVKD 480
QY 439 IPTLTGPGTLKRVYFMHLEKADPTKGRNEANVDNGHTDDQTPPVEVYVYTGKRPVKD 478
DB 439 IPTLTGPGTLKRVYFMHLEKADPTKGRNEANVDNGHTDDQTPPVEVYVYTGKRPVKD 478
QY 481 GGVTVQALAGASFVVRDONSPTANYLKI DETTKATVYKTAEBATFTTTADGLVDING 540
DB 479 GGVTVQALAGASFVVRDONSPTANYLKI DETTKATVYKTAEBATFTTTADGLVDING 538
QY 541 LKGYTYLEETVAPDDVYLLTNRIFVNEQSYGTENLVSEKVPNKKHGTLPSTGKG 600
DB 539 LKGYTYLEETVAPDDVYLLTNRIFVNEQSYGTENLVSEKVPNKKHGTLPSTGKG 598
QY 601 IYVYLSGAVLLILAGVTPARRRKENA 627
DB 599 IYVYLSGAVLLILAGVTPARRRKENA 625
RESULT 9
AAV00119
ID AAV00119 standard; protein; 560 AA.
XX
XX AC AAV00119;
XX
XX DT 20-APR-1999 (first entry)
XX
XX DE Enterococcus faecalis antigenic polypeptide fragment EP058.
XX
XX KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
XX KW detection; attenuation; antigenic.
XX
XX OS Enterococcus faecalis.
XX
XX PN WO9850554-A2.
XX
XX PD 12-NOV-1998.
XX
XX PF 04-MAY-1998; 98WO-US008959.
XX
XX PR 06-MAY-1997; 97US-0044031P.
XX PR 16-MAY-1997; 97US-004655P.

PR 14-NOV-1997; 97US-0066009P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Kunsch CA, Choi GH, Bailey C, Hromockyj A;
XX WPI; 1999-070095/06.
XX N-PSDB; AAX20109.
DR New isolated Enterococcus faecalis polynucleotides - used to develop
XX products for the detection of Enterococcus and for use in vaccines for
XX prevention or attenuation of Enterococcus infection.
PS Claim 9; Page 134; 301pp; English.
XX The present sequence represents an antigenic polypeptide fragment
CC isolated from Enterococcus faecalis. The present invention describes
CC genes, proteins and antigenic polypeptides isolated from E. faecalis. The
CC proteins can be used in vaccines for preventing or attenuating an
CC infection caused by a member of the Enterococcus genus in an animal. They
CC can also be used for detecting Enterococcus antibodies in a sample. The
CC nucleotide sequences can be used for detecting Enterococcus nucleic
CC acids. Products from the present invention can also be used for screening
CC compounds to identify agonists and antagonists of E. faecalis protein
CC activity
XX
SQ Sequence 560 AA;
Query Match 89.7%; Score 2892; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.4e-176;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 BENGSAQLVTHKKKMTDLPDLIONSGKEMSEFDKYQGLADVTFSYNTVEFYEQRA 92
DB 1 BENGSAQLVTHKKKMTDLPDLIONSGKEMSEFDKYQGLADVTFSYNTVEFYEQRA 60
QY 93 GASVDAAKQAVOSLTGPKRPAQGTDDANGNTVQLPKQNGDAVYTIKEPKEGVNAAT 152
DB 61 GASVDAAKQAVOSLTGPKRPAQGTDDANGNTVQLPKQNGDAVYTIKEPKEGVNAAT 120
QY 153 NMVVAFPVYEMTKQTDGSKYKGTBEIAVVAHIYPKVNVANDGSLHAKVGTAEENGANGAE 212
DB 121 NMVVAFPVYEMTKQTDGSKYKGTBEIAVVAHIYPKVNVANDGSLHAKVGTAEENGANGAE 180
QY 213 FVTSBSGSPGVTKYIQGVKGLVYMTTDKEQAKRFTGSKSYEIGENDFTEAENGTELT 272
DB 181 FVTSBSGSPGVTKYIQGVKGLVYMTTDKEQAKRFTGSKSYEIGENDFTEAENGTELT 240
QY 273 VKNLEVGSIYIIEEVAPNNAELIENQTKPTTIEANNQTPVEKTVKNDTSKYDKTTPSLD 332
DB 241 VKNLEVGSIYIIEEVAPNNAELIENQTKPTTIEANNQTPVEKTVKNDTSKYDKTTPSLD 300
QY 333 GKDVAIIGEKIKYQISVNIPLGIADKEGDANKYVKFNLDKHDALTFPDNVTSGEYAVYALY 392
DB 301 GKDVAIIGEKIKYQISVNIPLGIADKEGDANKYVKFNLDKHDALTFPDNVTSGEYAVYALY 360
QY 393 DGDVTIAPENYQVTEQANGFTVAVNPAYIPTLTPGGTLKFVYFMHLEKADPTKGPKNQA 452
DB 361 DGDVTIAPENYQVTEQANGFTVAVNPAYIPTLTPGGTLKFVYFMHLEKADPTKGPKNQA 420
QY 453 NVDNHGTDDQTPPTVEVVTGGRFKIKVDQDVATQALAGASFVVRDQNSDTANYLKIDBT 512
DB 421 NVDNHGTDDQTPPTVEVVTGGRFKIKVDQDVATQALAGASFVVRDQNSDTANYLKIDBT 480
QY 513 TGAATVWTKKAEATFTTTADGLVDITGKYGTYIIEEVAAPDDYLLTNRIEFVNNES 572
DB 481 TGAATVWTKKAEATFTTTADGLVDITGKYGTYIIEEVAAPDDYLLTNRIEFVNNES 540
QY 573 YGTENLVSPKVPNKHGT 592
DB 541 YGTENLVSPKVPNKHGT 560

RESULT 10
ABP43338
ID ABP43338 standard; protein; 560 AA.
XX AC ABP43338;
XX 05-AUG-2002 (first entry)
XX DE E faecalis EF058 antigenic fragment.
XX KW Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.
XX OS Enterococcus faecalis.
XX PN US2002045737-A1.
XX PD 18-APR-2002.
XX PF 04-MAY-1998; 98US-00071035.
XX PR 04-MAY-1998; 98US-00071035.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX WPI; 2002-425450/45.
XX DR N-PSDB; ABR98094.
XX PT New genes and polypeptides from Enterococcus faecalis, useful as vaccines
XX for preventing, treating or attenuating an infection caused by a member
XX of the Enterococcus genus in an animal, particularly E. faecalis.
XX PS Claim 9; Page 94; 255pp; English.
XX CC The present invention provides the protein and coding sequences of a
XX number of polypeptides from Enterococcus faecalis. The proteins can be
XX used as vaccines for preventing or attenuating an infection caused by a
XX member of the Enterococcus genus in an animal, particularly E. faecalis.
XX CC The polynucleotide is also useful for preventing or treating E. faecalis
XX infection. The present sequence is a protein of the invention
XX
SQ Sequence 560 AA;
Query Match 89.7%; Score 2892; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.4e-176;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 BENGSAQLVTHKKKMTDLPDLIONSGKEMSEFDKYQGLADVTFSYNTVEFYEQRA 92
DB 1 BENGSAQLVTHKKKMTDLPDLIONSGKEMSEFDKYQGLADVTFSYNTVEFYEQRA 60
QY 93 GASVDAAKQAVOSLTGPKRPAQGTDDANGNTVQLPKQNGDAVYTIKEPKEGVNAAT 152
DB 61 GASVDAAKQAVOSLTGPKRPAQGTDDANGNTVQLPKQNGDAVYTIKEPKEGVNAAT 120
QY 153 NMVVAFPVYEMTKQTDGSKYKGTBEIAVVAHIYPKVNVANDGSLHAKVGTAEENGANGAE 212
DB 121 NMVVAFPVYEMTKQTDGSKYKGTBEIAVVAHIYPKVNVANDGSLHAKVGTAEENGANGAE 180
QY 213 FVTSBSGSPGVTKYIQGVKGLVYMTTDKEQAKRFTGSKSYEIGENDFTEAENGTELT 272
DB 181 FVTSBSGSPGVTKYIQGVKGLVYMTTDKEQAKRFTGSKSYEIGENDFTEAENGTELT 240
QY 273 VKNLEVGSIYIIEEVAPNNAELIENQTKPTTIEANNQTPVEKTVKNDTSKYDKTTPSLD 332
DB 241 VKNLEVGSIYIIEEVAPNNAELIENQTKPTTIEANNQTPVEKTVKNDTSKYDKTTPSLD 300
QY 333 GKDVAIIGEKIKYQISVNIPLGIADKEGDANKYVKFNLDKHDALTFPDNVTSGEYAVYALY 392
DB 301 GKDVAIIGEKIKYQISVNIPLGIADKEGDANKYVKFNLDKHDALTFPDNVTSGEYAVYALY 360
QY 393 DGDVTIAPENYQVTEQANGFTVAVNPAYIPTLTPGGTLKFVYFMHLEKADPTKGPKNQA 452

DB 361 DGDVIAIENYQVTEQANGFTVAANPAYIPILTPGTLKFVYFHLNEKADPTKGFKQEA 420
QY 453 NVDNNGHTDDQPPPYEVVVTGGRFRIKVDGDTATATAGASFVVDQNSDTANYLKIDBT 512
DB 421 NVDNNGHTDDQPPPYEVVVTGGRFRIKVDGDTATATAGASFVVDQNSDTANYLKIDBT 480
QY 513 TKAATWVTKAEATFTTTADGLVDITGLKGTYYLLEETVAPDDVYLLTNRIEFPVNEGS 572
DB 481 TKAATWVTKAEATFTTTADGLVDITGLKGTYYLLEETVAPDDVYLLTNRIEFPVNEGS 540
QY 573 YGTENLVSPKVPKHKGT 592
DB 541 YGTENLVSPKVPKHKGT 560

RESULT 11
ABU8366
ID ABU8366 standard; protein; 560 AA.
XX
AC ABU8366;
XX
DT 07-JUL-2003 (first entry)
XX
XX
DE E. faecalis novel protein #110.
XX
XX Vaccine; endocarditis; bacteraemia; urinary tract infection; UTI;
KW intraabdominal infection; soft tissue infection; neonatal sepsis.
XX
OS Enterococcus faecalis.
XX
PN US2003017495-A1.
XX
XX 23-JAN-2003.
PD
PF 29-JUL-2002; 2002US-00206576.
XX
XX 06-MAY-1997; 97US-0044031P.
PR 16-MAY-1997; 97US-0046555P.
PR 14-NOV-1997; 97US-0066009P.
PR 04-MAY-1998; 98US-00071035.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
E1 Chol GH, Bailey C, Hromocky A, Kunach CA;
XX
XX WPI; 2003-416890/39.
DR N-Psdb; ACAB8058.
XX
XX New nucleic acid molecules and polypeptides from Enterococcus faecalis,
PT useful as vaccines for preventing or attenuating an enterococcal
PT infection in an animal, or for identifying Enterococcus faecalis in
PT biological samples.
XX
PS Claim 12; Page: 40pp; English.

CC The invention relates to a new isolated nucleic acid molecule comprising
CC a polynucleotide isolated from Enterococcus faecalis appearing as
CC ACAB87949-ACAB8196 (or sequences complementary to them or 95% identical to
CC them). Also included are the proteins encoded by the above nucleic acids,
CC making a recombinant vector (comprising inserting the isolated nucleic
CC acid molecule cited above into a vector), a host cell comprising the
CC vector (used to produce the protein), an isolated antibody specific for
CC the polypeptides, a hybridoma that produces the antibody, an isolated
CC polypeptide antigen comprising an amino acid sequence of an Enterococcus
CC faecalis epitope listed in the specification, a vaccine comprising one or
CC more E. faecalis polypeptides (and a pharmaceutical diluent, carrier or
CC excipient) where the polypeptides elicits protective antibodies in an
CC animal to a member of the genus Enterococcus; preventing or attenuating
CC an infection caused by a member of the genus Enterococcus in an animal
CC comprising administering to the animal the polypeptide and detecting
CC Enterococcus nucleic acids in a biological sample. The E. faecalis
CC nucleic acid molecules and polypeptides are useful as vaccines for

CC preventing or attenuating an enterococcal infection in an animal (e.g.
CC endocarditis, bacteraemia, urinary tract infection (UTI), intraabdominal
CC infection, soft tissue infection and neonatal sepsis). The polypeptides
CC are also useful for detecting Enterococcus aureus in immunoassays, as
CC epitope tags, as molecular weight markers, or for generating antibodies
CC that specifically bind E. faecalis polypeptides. The nucleic acid
CC molecules are also useful as probes for gene mapping, or for identifying
CC E. faecalis in biological samples. The kit and methods are useful for
CC detecting Enterococcus antibodies or nucleic acid molecules in a
CC biological sample. The present sequence is a novel E. faecalis
CC polypeptide of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030017495
XX
SQ Sequence 560 AA;
XX
Query Match 89.7%; Score 2892; DB 6; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.4e-176;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 BENGSAQLVIHKKMTDLPPLIIONSGKSEFPKYGGLADVTSTYVNTNEFPBORAA 92
DB 1 BENGSAQLVIHKKMTDLPPLIIONSGKSEFPKYGGLADVTSTYVNTNEFPBORAA 60
QY 93 GASVDAAKAQAOSLTPRGKPVACGTTDANGNVTVOLPKKONGDAVYTIKEEPKGVNAT 152
DB 61 GASVDAAKAQAOSLTPRGKPVACGTTDANGNVTVOLPKKONGDAVYTIKEEPKGVNAT 120
QY 153 NMVAFPVYEMIKOTDGSYKTYGTEBELAVVHIYPKNVVANDGSLHVKVYGTANEGLNGAE 212
DB 121 NMVAFPVYEMIKOTDGSYKTYGTEBELAVVHIYPKNVVANDGSLHVKVYGTANEGLNGAE 180
QY 213 FVTSKBSGSGTGYKTYGKQVKGGLYTTTDDKEQAKFTIGKSYEISENDFTEANGTGELT 272
DB 181 FVTSKBSGSGTGYKTYGKQVKGGLYTTTDDKEQAKFTIGKSYEISENDFTEANGTGELT 240
QY 273 VKNLEVGSYTILEBVAPNNALIEHQKTPPTIEANNQTPVEKTYKNDTSKYDKTTPSLD 332
DB 241 VKNLEVGSYTILEBVAPNNALIEHQKTPPTIEANNQTPVEKTYKNDTSKYDKTTPSLD 300
QY 333 GKDVAGIKKIKYQISVNIPLGIADKSGDANKYKKNLVKDKDAALTFPDVNTSGEYAVLY 392
DB 301 GKDVAGIKKIKYQISVNIPLGIADKSGDANKYKKNLVKDKDAALTFPDVNTSGEYAVLY 360
QY 393 DGDVIAIENYQVTEQANGFTVAANPAYIPILTPGTLKFVYFHLNEKADPTKGFKQEA 452
DB 361 DGDVIAIENYQVTEQANGFTVAANPAYIPILTPGTLKFVYFHLNEKADPTKGFKQEA 420
QY 453 NVDNNGHTDDQPPPYEVVVTGGRFRIKVDGDTATATAGASFVVDQNSDTANYLKIDBT 512
DB 421 NVDNNGHTDDQPPPYEVVVTGGRFRIKVDGDTATATAGASFVVDQNSDTANYLKIDBT 480
QY 513 TKAATWVTKAEATFTTTADGLVDITGLKGTYYLLEETVAPDDVYLLTNRIEFPVNEGS 572
DB 481 TKAATWVTKAEATFTTTADGLVDITGLKGTYYLLEETVAPDDVYLLTNRIEFPVNEGS 540
QY 573 YGTENLVSPKVPKHKGT 592
DB 541 YGTENLVSPKVPKHKGT 560

RESULT 12
ABU13617
ID ABU13617 standard; protein; 560 AA.
XX
AC ABU13617;
XX
DT 26-FEB-2003 (first entry)
XX
XX Enterococcus faecalis EF040 polypeptide #110.
DE Enterococcus faecalis EF040 polypeptide #110.
XX
KW EF040; immunostimulant; antibacterial; gene mapping.

XX Enterococcus faecalis.
 OS
 XX US6448043-B1.
 PN
 XX 10-SEP-2002.
 PD
 XX 04-MAY-1998; 98US-00071035.
 PF
 XX 06-MAY-1997; 97US-0044031P.
 PR 16-MAY-1997; 97US-004655P.
 PR 14-NOV-1997; 97US-0066009P.
 PR 14-NOV-1997; 97US-0066009P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Choi GH, Bailey C, Hromockyj A, Kunsch CA,
 PI
 DR WPI; 2003-089120/08.
 DR N-PSDB; ABX61664.
 XX
 XX New EF040 polypeptides and polynucleotides from Enterococcus faecalis,
 PT useful for generating an immune response against E. faecalis and other
 PT Enterococcus species, and as vaccines against other bacterial genera.
 XX
 PS Example 1; Col 115-116; 146pp; English.

CC The invention relates to polynucleotide fragments of a gene from
 CC Enterococcus faecalis, EF040, and the polypeptides encoded by them. The
 CC polypeptides are useful in detecting E. faecalis, as epitope tags, as
 CC molecular weight markers on SDS-PAGE gels or for molecular sieve gel
 CC filtration columns, in generating antibodies that specifically bind to
 CC the E. faecalis polypeptides, in generating an immune response against E.
 CC faecalis and other Enterococcus species and as vaccines against other
 CC bacterial genera. The polynucleotides are useful as probes for gene
 CC mapping and for identifying E. faecalis in biological samples. Sequences
 CC ABU13508-ABU13755 represent EF040 polypeptides of the invention. Note:
 CC The sequence data for this patent can also be obtained from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 XX

Sequence 560 AA;

Query Match 89.7%; Score 2892; DB 6; Length 560;
 Best Local Similarity 100.0%; Pred. No. 1.4e-176;
 Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 EENGSAQVLVHKKKMTDLPDLIONSCKEMSEFDPKYGGLADVTFSYNNVTFEYQRA 92
 DB 1 EENGSAQVLVHKKKMTDLPDLIONSCKEMSEFDPKYGGLADVTFSYNNVTFEYQRA 60
 QY 93 GASVBAKQAVOSLTPGKPVAGTTDANGNVTVOLPKONGKDAVYTIKEPKEGVVAAT 152
 DB 61 GASVBAKQAVOSLTPGKPVAGTTDANGNVTVOLPKONGKDAVYTIKEPKEGVVAAT 120
 QY 153 NMVAVFPVYEMIKQTDGSIKYGTEBELAVVHIYPRKVVANDSGLHYKVGTAENEGINGAE 212
 DB 121 NMVAVFPVYEMIKQTDGSIKYGTEBELAVVHIYPRKVVANDSGLHYKVGTAENEGINGAE 180
 QY 213 PFISSBSGPGTVKXIQGVKGLVYMTTDEKQAKPTIKSKYEIGENDTEAENGNGELT 272
 DB 161 PFISSBSGPGTVKXIQGVKGLVYMTTDEKQAKPTIKSKYEIGENDTEAENGNGELT 240
 QY 273 VGNLEVGSIYILEEVAAPNNAELIENQKPTTEANNQTPVEKTVANDSKDKTTPSLD 332
 DB 241 VGNLEVGSIYILEEVAAPNNAELIENQKPTTEANNQTPVEKTVANDSKDKTTPSLD 300
 QY 333 GKDVAIGEKIKYQISVNIPLGIADKEGANKYVKFNLVDKHDALTFDNTSGEYAVYALY 392
 DB 301 GKDVAIGEKIKYQISVNIPLGIADKEGANKYVKFNLVDKHDALTFDNTSGEYAVYALY 360
 QY 333 DDDTVIAPENVQVTOANGFTVAVNPAYIPTLTPGGTLKFKVYFMHLNEKADPTKGRKNA 452
 DB 361 DDDTVIAPENVQVTOANGFTVAVNPAYIPTLTPGGTLKFKVYFMHLNEKADPTKGRKNA 420

QY 453 NVDNHGTDDQTPPEVEVTVGKRFYKVDGVTATQALAGASFVVRDQNSDPTANYKIDET 512
 DB 421 NVDNHGTDDQTPPEVEVTVGKRFYKVDGVTATQALAGASFVVRDQNSDPTANYKIDET 480
 QY 513 TKAATWVKTKABATFTTTADGLVDITGLKRGTYYLEEFTVAPDDYVLLTNRIEFVNEQS 572
 DB 481 TKAATWVKTKABATFTTTADGLVDITGLKRGTYYLEEFTVAPDDYVLLTNRIEFVNEQS 540
 QY 573 YGTENLVSPKVPKHKGT 592
 DB 541 YGTENLVSPKVPKHKGT 560

RESULT 13

ID ADY38954 standard; protein; 560 AA.
 XX ADY38954;
 AC
 XX
 XX 05-MAY-2005 (first entry)
 DT
 XX
 XX DE Novel Enterococcus faecalis protein sequence SeqID220.
 XX
 XX protein purification; DNA purification; antibacterial; vaccine;
 KW enterococcus faecalis infection.
 XX
 XX

OS Enterococcus faecalis.

PN US2005043528-A1.

PD 24-FEB-2005.

PF 06-AUG-2004; 2004US-00912362.

PR 06-MAY-1997; 97US-0044031P.

PR 16-MAY-1997; 97US-004655P.

PR 14-NOV-1997; 97US-0066009P.

PR 04-MAY-1998; 98US-00071035.

PR 29-JUL-2002; 2002US-00206576.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Choi GH, Bailey C, Hromockyj A, Kunsch CA;

PI WPI; 2005-180840/19.

DR N-PSDB; ADY38953.

PT New isolated Enterococcus faecalis nucleic acid useful for diagnosing,

PT preventing and/or attenuating infection by Enterococcus faecalis in

PT animals, particularly in humans.

PS Claim 9; SEQ ID NO 220; 28pp; English.

CC This invention relates to novel genes and the proteins they encode
 CC isolated from Enterococcus faecalis. The invention may be useful for the
 CC development of compounds with an antibacterial activity or a vaccine. The
 CC present invention is useful for the diagnosis, prevention and/or
 CC attenuation of infection by Enterococcus faecalis in animals.
 CC particularly in humans. The present sequence is that of a novel
 CC Enterococcus faecalis protein of the invention. Note: The present
 CC sequence does not form part of the printed specification but was obtained
 CC in electronic form from the USPTO web.

Sequence 560 AA;

Query Match 89.7%; Score 2892; DB 9; Length 560;
 Best Local Similarity 100.0%; Pred. No. 1.4e-176;
 Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 EENGSAQVLVHKKKMTDLPDLIONSCKEMSEFDPKYGGLADVTFSYNNVTFEYQRA 92
 DB 1 EENGSAQVLVHKKKMTDLPDLIONSCKEMSEFDPKYGGLADVTFSYNNVTFEYQRA 60

QY 93 GASVDAKQAVOSLTPGKPAOCTTDANGNVTYQLPKONGDAVYTIKEEPEGVAAAT 152
DB 61 GASVDAKQAVOSLTPGKPAOCTTDANGNVTYQLPKONGDAVYTIKEEPEGVAAAT 120
QY 153 NNVAAPVYEMIKQTDGSKYKGTSEELAVVHIYKPNVANDGSLHVKVGTAEENGANGAE 212
DB 121 NNVAAPVYEMIKQTDGSKYKGTSEELAVVHIYKPNVANDGSLHVKVGTAEENGANGAE 180
QY 213 FVYSSESGPVTYKIQGVKDGLYTTTDEKQAKPTGSKYSYIEGNDPTEANGTEGELT 272
DB 181 FVYSSESGPVTYKIQGVKDGLYTTTDEKQAKPTGSKYSYIEGNDPTEANGTEGELT 240
QY 273 VKNLEVGSIYLBKVPNNNAELIENOTKPTTEANNQTPVEKTVNNDTSKYDKTTPSLD 332
DB 241 VKNLEVGSIYLBKVPNNNAELIENOTKPTTEANNQTPVEKTVNNDTSKYDKTTPSLD 300
QY 333 GKDVAIGEKIKYQISVNIPLGIADKEGDANKYKFNLYDKHDAALTFDNTSGEYAYALY 392
DB 301 GKDVAIGEKIKYQISVNIPLGIADKEGDANKYKFNLYDKHDAALTFDNTSGEYAYALY 360
QY 393 DDDTVIAPENYQVTEQANGFTVAVNPAYIPPLTPGGTLKPYVPMHLEKADPTKGFKNBA 452
DB 361 DDDTVIAPENYQVTEQANGFTVAVNPAYIPPLTPGGTLKPYVPMHLEKADPTKGFKNBA 420
QY 453 NVDNGHTDQTPPTVEVTVGSKRFKRVDDVATATQALAGSPVVRQNSDPTANYLKIDET 512
DB 421 NVDNGHTDQTPPTVEVTVGSKRFKRVDDVATATQALAGSPVVRQNSDPTANYLKIDET 480
QY 513 TKAATVTKKAEATFTTTADGLVDITGLKYGTYLEETVAPDDVYLTNRLEFVNEQS 572
DB 481 TKAATVTKKAEATFTTTADGLVDITGLKYGTYLEETVAPDDVYLTNRLEFVNEQS 540
QY 573 YGTENLVSPKVPNNKHKT 592
DB 541 YGTENLVSPKVPNNKHKT 560
RESULT 14
AD084895
ID AD084895 standard; protein; 625 AA.
AC AD084895;
XX
DT 29-JUL-2004 (first entry)
DE E faecium surface anchored LPXTG protein Segid19.
XX
KM LPXTG; cell wall-anchored surface protein; Gram positive bacterium;
KM extracellular matrix molecule; sequence database; C-terminal;
KM immunoglobulin-like fold region; Ig-like fold region; antibacterial;
KM vaccine; gene therapy; infection; medical device; prosthesis;
XX premature newborn; AIDS; debilitated cancer; bone marrow transplantation.
OS Enterococcus faecium.
XX
PN WO2004025416-A2.
XX
PD 25-MAR-2004.
XX
PF 15-SEP-2003; 2003WO-US028789.
XX
PR 13-SEP-2002; 2002US-0410303P.
XX
PA (TEXA) UNIV TEXAS A & M SYSTEM.
PA (INH-) INHIBITEX INC.
PA (UABR-) UAB RES FOUND.
XX
PI Hook M, Xu Y, Sillanpaa JV, Sthanam N, Ponnuraj K, Patti JM,
PI Hutchins JT, Hall A;
XX
DR WPI; 2004-315684/29.
DR N-PsDB; AD084894.

XX
PT Identifying LPXTG-containing cell wall-anchored surface proteins from
PT Gram positive bacteria, for treating infection caused by the bacteria,
PT comprises searching sequence information database for the sequence having
PT LPXTG-motif.
XX
PS Claim 16; SEQ ID NO 19; 96pp; English.
XX
CC This invention relates to a novel method of identifying LPXTG-containing
CC cell wall-anchored surface proteins from Gram positive bacteria that bind
CC to an extracellular matrix molecule which comprises searching a database
CC of sequence information for a putative protein sequence having the LPXTG-
CC motif in its C-terminal region and analyzing the sequence for the
CC presence of one or more immunoglobulin (Ig)-like fold regions. The
CC invention may be useful for the production of compounds with an
CC antibacterial activity or for production of a vaccine. In addition the
CC disclosed sequences may be useful for gene therapy. The antibody is
CC useful for treating or preventing an infection of Gram-positive bacteria
CC in a human or animal patient. The method and the proteins are useful in
CC generating antibodies for treating and preventing the spread of
CC infections of Gram positive bacteria, for interfering with, or inhibiting
CC binding interactions by Gram positive bacteria, for monitoring the level
CC of gram positive bacterial antigens, or antibodies recognizing the
CC antigens in a human or animal patient suspected of containing the
CC antigens or antibodies, in preventing or reducing infection of medical
CC devices and prostheses caused by such organisms, and in treating or
CC preventing infections in highly susceptible groups such as premature
CC newborns, AIDS and debilitated cancer patients, and bone marrow
CC transplantation. The present sequence is that of a surface anchored LPXTG
CC protein identified using the method of the invention.
XX
SQ Sequence 625 AA;
Query Match 75.8%; Score 2446; DB 8; Length 625;
Best Local Similarity 74.2%; Pred. No. 5,9e-148;
Matches 462; Conservative 64; Mismatches 95; Indels 2; Gaps 1;
QY 1 MKQKKWYVSTLLILPLFTSVLGTTTAFKENGESAKVYHKKKMTDLPDPLIONSG 60
DB 1 MKQKKWYVSTLLILPLFTSVLGTTTAFKENGESAKVYHKKKMTDLPDPLIONSG 58
QY 61 KEMSEFDKQGLADYFYSIVYNTNEFYEDQRAAGASVDAKQAVOSLTPGKPAOCTTDAN 120
DB 59 KEMSEFDKQGLADYFYSIVYNTNEFYEDQRAAGASVDAKQAVOSLTPGKPAOCTTDAN 118
QY 121 GNVTVQLPKQNGKQAVYTIKEEPEGVAAATNNVAPVYEMIKQTDGSKYKGTSEELAV 180
DB 119 GNVTVQLPKQNGKQAVYTIKEEPEGVAAATNNVAPVYEMIKQTDGSKYKGTSEELAV 178
QY 181 VHIYKPNVANDGSLHVKVGTAEENGANGAEFVYSKESGSPGTYKIQGVKDGLYTTWT 240
DB 179 VHIYKPNVANDGSLHVKVGTAEENGANGAEFVYSKESGSPGTYKIQGVKDGLYTTWT 238
QY 241 DKEQAKPTGSKYSYIEGNDPTEANGTEGELTVKNLEVGSIYLBKVPNNNAELIENOTK 300
DB 239 DKEQAKPTGSKYSYIEGNDPTEANGTEGELTVKNLEVGSIYLBKVPNNNAELIENOTK 298
QY 301 TPFTTEANNQTPVEKTVNNDTSKYDKTTPSLDGLKQVAIGEKIKYQISVNIPLGIADKEGD 360
DB 299 TPFTTEANNQTPVEKTVNNDTSKYDKTTPSLDGLKQVAIGEKIKYQISVNIPLGIADKEGD 358
QY 361 ANKYKVPNLYDKHDAALTFDNTSGEYAYALYDGPVTVAPENYQVTEQANGFTVAVNPAY 420
DB 359 ANKYKVPNLYDKHDAALTFDNTSGEYAYALYDGPVTVAPENYQVTEQANGFTVAVNPAY 418
QY 421 IPTLTPGGTLKPYVPMHLEKADPTKGFKNBAVNDNGHTDQTPPTVEVTVGSKRFKRV 480
DB 419 IPTLTPGGTLKPYVPMHLEKADPTKGFKNBAVNDNGHTDQTPPTVEVTVGSKRFKRV 478
QY 481 GPDVATQALAGSPVVRQNSDPTANYLKIDETTKAATVTKKAEATFTTTADGLVDITG 540
DB 479 GPDVATQALAGSPVVRQNSDPTANYLKIDETTKAATVTKKAEATFTTTADGLVDITG 538


```
QY 541 LKGYTYLEETVAPDDVYLLTNRIEFVNVESGYTTEMLVSPKYPNKHGKTLPTSGKG 600
DB 539 LKGYTYLEETVAPDDVYLLTNRIEFVNVESGYTTEMLVSPKYPNKHGKTLPTSGKG 598
QY 601 IYVYLGSGAVLLLLIAGVYFARRR 623
DB 599 IYVYLGSGAVLLLLIAGVYFARRR 621

RESULT 15
ADV16668
ID ADV16668 standard; protein; 625 AA.
XX
AC ADV16668;
XX
DT 24-FEB-2005 (first entry)
XX
DE E. faecium hyperimmune serum reactive antigen protein - SEQ ID 363.
XX
KW antigen; antibacterial; vaccine; enterococcus infection; infection;
KW pharyngitis; impetigo; rheumatic fever; antipyretic; antirheumatic;
KW immunosuppressive; sepsis.
XX
OS Enterococcus faecium.
XX
PN WO2004106367-A2.
XX
PD 09-DEC-2004.
XX
PF 26-MAY-2004; 2004WO-EP005664.
XX
PR 30-MAY-2003; 2003EP-00450137.
XX
PA (INTE-) INTERCELL AG.
PI Meinke A, Nagy E, Hanner M, Gelbmann D;
XX
DR MPI; 2005-039707/04.
XX
N-PSDB; ADV16652.
XX
PT Novel isolated nucleic acid molecule encoding hyperimmune serum reactive
PT antigen e.g., EF0020, EF0032, EF0062, EF0149, EF0196, EF0253, EF0270,
PT EF0298 or its fragment, useful for producing vaccine against enterococcal
PT infection.
XX
PS Example 3; SEQ ID NO 363; 175pp; English.
XX
CC The invention relates to a novel isolated nucleic acid molecule encoding
CC a hyperimmune serum reactive antigen e.g. SEQ ID 171-340, 357-372 or 425-
CC 476, or its fragment. The antigen of the invention demonstrates
CC antibacterial activities and may be useful for producing a pharmaceutical
CC preparation, particularly a vaccine, against enterococcal infection.
CC Conditions associated with bacterial infection which may be prevented or
CC treated include bacterial pharyngitis, scarlet fever, impetigo, rheumatic
CC fever, necrotizing fasciitis and sepsis in humans. The current sequence
CC is that of an Enterococcus faecium hyperimmune serum reactive antigen
CC protein of the invention.
XX
SQ Sequence 625 AA;

Query Match 75.8%; Score 2446; DB 9; Length 625;
Best Local Similarity 74.2%; Pred. No. 5.9e-148;
Matches 462; Conservative 64; Mismatches 95; Indels 2; Gaps 1;

QY 1 MKQULKWVTVSTLLIPLFTSVLGTTTAPAEENGESQVLIHKKKMTLDPDPLIUNG 60
DB 1 MNHKKINVMGLVPLIPLTNSFGAKKFAEET--AAQVILHKKKMTLDPDPLIUNG 58
QY 61 KEMSEFDKQGLADYTESLYNTNFEYEBORAAASVDAKQAVOSLTPGKPVAAOGTTDAN 120
DB 59 KEMSEFDQYQGLADISFSYNTVTOEFYARQDKASVDAKQAVOSLTPGTPVAAOGTTDAD 118
QY 121 GNVTVQLPKKQNGKDAVYTIKEPEKGVAAATNMVAVFVEMIKOTDGSYKYGTTELAV 180
```

```
DB 119 GNVTVLSLPKQNGKQDAVYTIKEPEKDGVSAAANMVLAPFVYEMIKQADGSYKYGTTELD 178
QY 181 VHTIPKXNVANDGSLHVKKQVTAENGLNGAEFVYSKESGPGVYKIQYKVDGLYTWTT 240
DB 179 IHLVPKNTVAGNDGTLKVKIKIETAENBALNGAEFIIKSKEGTPSVKRYIQSVTDGLYTWTT 238
QY 241 DKEAKRFITGKSVEIGENDPTEBAENGTELTVKLVEGYSYLIEVKAPNNNAELIENQTK 300
DB 239 DQTKAKKHITGHSIDIGANNDAEASIEKGQILVNHLEVGKXNLEVKAPDNNEMIEKQTI 298
QY 301 TPTTIEANNQTPVEKTVKNDTSKYDKTTPSLDGKQVAIGEKIKYQISVNIPLGIADKEGD 360
DB 299 TPTTEIANSQTPVEKTIKNDTSKYDKTTPQNGKQVAIGEKIYQISVNIPLGIADKEGT 358
QY 361 ANKYKFEVLVDKHDAAALTFDNTSGEYAYALYDGTVAAPENYQVTEQDNGFTVAVNAY 420
DB 359 QNKYTFPLIDTHDAALTFDNDSSGTAYALYDGNKEIDPVNYSVTEQDNGFTVAVDNY 418
QY 421 IPTLPGGTLKPVYPMHINEKADPTKGFKNENAVNDGHTDQTPTEVEVTGKRFIYD 480
DB 419 IPSLTPGGTLKPVYPMHINEKADPTKGFKNENAVNDGHTDQTPTEVEVTGKRFIYD 478
QY 481 GDVTAQALAGASFVVRDQNSDGTANYLKI DETTKAATVKTAKAETTTTADGLVDITG 540
DB 479 GDVTSIDQTLAGAEFVVRDQSDGTAKYLSIDSTKAVSWSAKESATVFTTISNGLI DVTG 538
QY 541 LKGYTYLEETVAPDDVYLLTNRIEFVNVESGYTTEMLVSPKYPNKHGKTLPTSGKG 600
DB 539 LKGYTYLEETVAPDDVYLLTNRIEFVNVESGYTTEMLVSPKYPNKHGKTLPTSGKG 598
QY 601 IYVYLGSGAVLLLLIAGVYFARRR 623
DB 599 IYVYLGSGAVLLLLIAGVYFARRR 621
```

Search completed: July 6, 2006, 02:31:26
Job time : 201 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 02:31:46 ; Search time 42 Seconds
(without alignments)

1436.379 Million cell updates/sec

Title: US-10-661-809A-13

Perfect score: 3225

Sequence: 1 MKQLKKVWYVSTLLILPL.....GAVLLIAGVYFARRRKENA 627

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	458	14.2	507	2	SS2348
2	296.5	9.2	665	2	P95053
3	296	9.2	1530	2	AB1396
4	267	8.3	1612	2	AB1347
5	250.5	7.8	1622	2	AB1717
6	247	7.7	1806	2	AP1717
7	244.5	7.6	534	2	AJ3347
8	242.5	7.5	793	2	AB1094
9	234.5	7.3	1816	2	P83901
10	229	7.1	681	2	AC1458
11	228	7.1	1983	2	G86643
12	220.5	6.8	537	2	A35400
13	216.5	6.7	1661	2	A83695
14	215	6.7	614	2	P86719
15	212.5	6.6	940	2	AB1744
16	202.5	6.3	393	2	G95053
17	196.5	6.1	533	2	A35259
18	195.5	6.1	1873	2	T30944
19	193.5	6.0	940	2	AD1374
20	192.5	6.0	1039	2	T30856
21	190.5	5.9	586	2	AD1458
22	189	5.9	5291	2	P90696
23	186.5	5.8	571	2	AI1094
24	185	5.7	2551	2	B98047
25	182	5.6	5188	2	B85547
26	180	5.6	745	2	S06992
27	180	5.6	745	2	B33856
28	178	5.5	893	2	B95053
29	177.5	5.5	1092	2	S42798

30	170	5.3	1939	2	D97316	probable S-layer p
31	169.5	5.3	1386	2	AC1533	surface protein (L
32	169.5	5.3	1417	2	H90670	probable invasin I
33	169.5	5.3	1417	2	D85521	probable adhesin e
34	169.5	5.3	4936	2	AH2515	hypothetical prote
35	168	5.2	13055	2	T16580	hypothetical prote
36	167	5.2	562	2	AE1335	probable peptidogl
37	166	5.1	2059	2	D82671	surface protein XF
38	165.5	5.1	814	2	I40048	S-layer protein pr
39	165.5	5.1	1185	2	A42404	collagen adhesin -
40	164.5	5.1	593	2	S00128	protein G precursor
41	164.5	5.1	1092	2	T30214	fibritogen-binding
42	164.5	5.1	1483	2	C97012	probably celluloso
43	164.5	5.1	1684	2	JM0057	gravin - human
44	164	5.1	1965	2	S75200	fat protein - Syne
45	161	5.0	664	2	S42574	streptococcal surf

ALIGNMENTS

RESULT 1

SS2348
hypothetical protein 2 - Lactobacillus leichmannii
C:Species: Lactobacillus leichmannii
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: SS2348
R:Schenk-Groeninger, R.
submitted to the EMBL Data Library, January 1995
A:Reference number: S52347
A:Accession: S52348
A:Molecule type: DNA
A:Residues: 1-507 <SCH>
A:Cross-references: UNIPROT:Q48707; UNIPARC:UP10000082PDE; EMBL:X01869; NID:G666067; PIL
A:Experimental source: DSM 20076

Query Match	14.2%; Score 458; DB 2; Length 507;
Best local similarity	24.9%; Pred. No. 5.6e-18;
Matches	166; Conservative 81; Mismatches 197; Indels 224; Gaps 26;
QY	11 VSTLLILPLFTSVLGT--TTAPAEENGSAQIVHKKKTDLDPDLIQSGKEMSFDR 69
DB	9 ITGLIMILPLILSLSPATQKVASADTNNSVKTLH-KRVFDSAQAKQVIGELMDFG- 66
QY	70 QGLADVTESIVYNTBFEQRAAGASVDAKAYOS-----LTPGKPAQGT--- 117
DB	67 TELNGVTFKAVYNTVDHLSLRKSGSDADAVTAIQSDAKSDNLPYAGSAIATETIAT 126
QY	118 -----DANGNTVQLPKQNGDAVYTIKEBPKEGVAAATNVVAPVYE 162
DB	127 KEDDGIAPNDNLMDSDGNYQTVL-----FVETDSPDTDTQQAAPVLTWPIY- 175
QY	163 MIKQTDGSKYKTEBELAVH---IYKQNVANDSLHVKVGTAEINEGALGAEFVYSKE 219
DB	176 --KTSPTS-----AIWHDIQIYKQV----- 194
QY	220 GSPGVYKYIGVADGLYTTTIDKEQAKRITTSKSYEIGENDPTEANGTELTVKNLBYG 279
DB	195 ----- 194
QY	280 SYLLEVKAPNNAEILNQTKPTTEIANNQTPVEKTVKNDYSKYDKTTPSDGKDV--- 336
DB	195 -----KSPITPKDL-DEASKDLAVTLPDGSTIYNA 224
QY	337 AIGEKIKQISVINPLGADKEGDANKYKFLVLDKHAALFPDNTVSEIYALVDSGT 396
DB	225 QYKSGFQVYITVNVPMNKKD-----TFNVVDPKPDGIDIDAST-----VSIDGLT 271
QY	397 VIAPENVQTEQANGPTVA--VNPAVIPTLTGSGTLKFYFPHLNEKADPTKGFKEA-- 452
DB	272 --KSTDIYVKKONGVQVFFKTTSAVQALA-GKSLITTYKATLTNNALPDDAIGNTATL 328
QY	453 NVDNGHTDQTPPT-VEVVTGKRFLKVDGVTATQALAGAP--VVRDQNSDTANYLKI 509

Db 329 SGNCTNTSTPANGPRITYTGAQFVKDSQ--SNKTLAAGFQVKNDSNGNYSY--- 383
QY 510 DETTKAA---TWTKKATATFTTTADGLVDITGKYG-----TYLEETVAPDDYV 558
Db 384 --ATQASQDSYWNMSATATYTTSDANGLVALKGLSYSDKLDGSEVALLEIQAPDGYA 441
QY 559 LITNRIEFPVNGSQYGTTENLVSPKVPKHKHGLTLPSTGKGIYVYIGSGAVLLILIA-GV 617
Db 442 KLDSPVKSTTQSGFSDSNKIT----IDNTKGLLPSTGKGKGIYFLAIGVIMVAFGG 497
QY 618 YEARRRKE 625
Db 498 YKAIKKHE 505

RESULT 2

P95053

cell wall surface anchor family protein [imported] - Streptococcus pneumoniae (strain TI
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: F95053
R:Jettelein, H.; Nelson, K.E.; Paulsen, I.T.; Bisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21537209; PMID:1146316
A:Accession: P95053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-665 <KUR-
A:Cross-references: UNIPROT:Q97S62; UNIPARC:UP1000005144B; GB:AE005672; PIDN:AAK74623.1;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0463

Query Match 9.2%; Score 296.5; DB 2; Length 665;
Best Local Similarity 23.0%; Pred. No. 7.2e-09; Indels 247; Gaps 38;
Matches 176; Conservative 84; Mismatches 259;
QY 1 MKOLKKWYTVSTLLILPLFTSVIGTTTAFBENGESAQVLVHKKMTDLPDPLIONSG 60
Db 1 MXSINKFLMLALILTA---SSLPASATVFA-AGTTTSVYHKLALIDGMDKIANE- 55
QY 61 KEMSEPDKIYQ-----LADVTSIYNVTNEFEQ--RAAGASVDAKAQVQSL 106
Db 56 --LETGVNAGKRVGVLPANAKEIAGVMEVMTNNEIIDENGQTLGAVNI DQTFELSGA 112
QY 107 TPGKVAOQTTPDANGV--TVQLPKKQNGKDAVYTIKEPKGSVAATMMVAAPYEMI 164
Db 113 MPAIMKRC-LTEABEGKENTANLP-----AAKTKIYEIH 145
QY 165 KQT---DGSYKGTBELAV-----VHIYPRNV-----VANDGSLH 196
Db 146 SLSTVYGEAGATLTGSKAPIRIELPLANDVAHVPRKTBKPKIDKFKGKANDPDR 205
QY 197 VKKGTAEVBGLANGAEF-VISKSBSFGTVKTIQGVKGLYVTTDDKEQAKPITGSKYE 255
Db 206 VKKDPVFNHQVDVVEYEIVTK--IPALANYATA-----SDRMTBELAV- 249
QY 256 IEDNPTEBANGTGLTVKN--LEVGSYLLEEVK-----APNNAELIENQGT 301
Db 250 -----NKGIYKTVTDVVALLEAGDYALTEVATGFDLKLTDAGLAKVNDQAEKTVKI 300
QY 302 PTTIANNQTPYEKTVKND-----TSKVDKTPSLDGKQDAVIGEKIKYQISV 348
Db 301 TYSATLNDRAIIVEPESNDVTFTYGNPNPHGNTPKENK--PNEG-DLTLTKTWDAQGA 357
QY 349 NPLPLGADGEGDANKVYKFNLDKIDALTFD--NVTSGEYVALYDQGTVAPEYVQYT 406
Db 358 PIPAG-----AEATFDLVNMQTKVVOQTV---TLTTDKN---T 389

QY 407 EQANGPTVAVNPAYIPTLIPGGLTKFYVFMILNEKADPTKGFKNENAVNDGHTDQTPPT 466
Db 390 VTVNLGDKNTEYKFFERSIKGSADYQELITTAGELIA--VKMKDE---NPKPLDPTIE- 442
QY 467 VEVVYGGKRFIKV-DGVYATQALAGAFVVDONSDBA-NYL--KIDE-----T 512
Db 443 -KVYVYGGKRFVANDKD-----NRLAGAEFVL--ANADNAGQYLAKKADKVSQEKQLVYT 495
QY 513 TK-----AATV--KTKAETTF 528
Db 496 TTDALDRAVVAAYNALTAQOQTQOEKEXVDKQAAYNAAVIANNAFEWVADNDENVYKL 555
QY 529 TTTADGLVDITGKGYTYLLEETVAPDDYVLLTNRIEFPVNGSQYGTTENLV----- 580
Db 556 VSDAQGRREITGLAGTYLLEETKQPAGYALLTSQKPEVNTSISATQGGILEYTAGSK 615
QY 581 -SPKVPKHKHGLTLPSTGKGIYVYIGSGAVLLIAGVYFARRRKE 625
Db 616 DATKVVNKR-KITIPQTGGIGTIIFAVAGAAIMGIAVAYVKNKD 660

RESULT 3

AH1396

peptidoglycan anchored protein (LPXTG motif) [imported] - Listeria monocytogenes (strain
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH1396
R:Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Biocker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fslh, H.
D.; Jones, L.M.; Karsic, U.
Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Schaefer, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voos, H.; Weiland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1396
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1530 <GLA>
A:Cross-references: UNIPROT:Q8Y479; UNIPARC:UP10000054CB4; GB:NC_003210; PIDN:CAD00654.1
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2576

Query Match 9.2%; Score 296; DB 2; Length 1530;
Best Local Similarity 22.2%; Pred. No. 2.6e-08; Indels 276; Gaps 30;
Matches 167; Conservative 71; Mismatches 239;
QY 28 TPAFBENGESAQVLVHKKMTDLPDPLIONSGKEMSEPDKIYQGLADVTFISIYNVTNEFY 87
Db 876 TTGGDTGCTGCTKIILNK---VDKADPSIPLEG-----ATFDLY----- 911
QY 88 ECPAAGASVDAKAQVOSITPGKVAOQTTPDANGVTVQLPKKQNGKDAV--YTIKE-E 143
Db 912 ---ANDEKVDQ-----TTDKNGIYEPD-----DLVYGDYTLKEVS 944
QY 144 PKEGVVAATNMVAFVYEMIKQDGSYKGTBELAVVHIYPRNVVANDGSLHVKVGYA 203
Db 945 ABEGYTLTPASTENIQV-----KLDEKRVYOVVMEKMPKETGEVHALVKTDXA 993
QY 204 ENEGANGAEFVYSKESGP-----GTVYKIQGVKGLYTW 238
Db 994 TQATLGAHFSYDYSGAELOGLTDDENGELTIHNLDSGYLLKETAPBEGYKLSKTM 1053
QY 239 TTDEQ-----AKRPITGKSYEIGENDFTFA 264
Db 1054 EFSVSGQVDALIEIOAENEKDLGEAVLTVVDSRTNAKLSGAFNLINLSGVEYIQTNLVSD 1113
QY 265 ENGGEELTVKNLEVGSYILEEVKAPNNAELINQTKPTPTIRANNQTPYEKTVKND-TSK 323
Db 1114 EN--GEIRVQNPEDDYAFQETEAFTPNYDLATN--TWPTIYVAGQTSATMTAENNKTK 1169

```
QY 324 VDKTPS--LDKQVAIGEKIKYQISVNIPLGIADKEGDANKYKFNLYDKHDAALTFDN 381
D 1170 PVDVIGSEVLIVKQDSATGETLEGAV---FDLMTRD-----GAIVASNLTTDANGELTYVN 1221
QY 382 VTSGEYVALYDGDVIAIPENYQVTEQANGFTVAANPAPIPLT-----PGGTLK 431
D 1222 LAFGRKYSF-----KETKABEGVELADVMEFTIAPNQPEKITTYLAENTKAPIPDAGSVK 1276
QY 432 FYVFMHLNKRADPTKPK-----K-----NE 451
D 1277 II-----KQSENGRILAGAEFSLIANGESTLQTNLKTDEAGELEVNNLAPGNRIOE 1329
QY 452 AAVNDGHTDDQTPPTVEV-----TGKRPFKVDGVTATQAL 489
D 1330 TRAPDQYQLESTWQGEIYVANDTSQVTVIAENAKLEBDVAETGAVALLIKTDSF-TGTR-L 1387
QY 490 AGASPVVRDQNSDTANYLKI DETTKAATWVKTAEMTPTTTADGLVDITGLKRYGYTLE 549
D 1388 SGAVFSLDDESG-----KVIQANLTTDENGEIPIDGLTPGNVSLK 1427
QY 550 ETVAPDDYVL-----LTKRIEFV--NEOSYGTENLV-----PR--KVPNK 588
D 1428 ETKAPDGYELAQPMNFQIVKQGVDAVILKAENSPILANGAISFEGDETDKPSISIPVR 1487
QY 589 HKGT-----LPSTGAKGY--VYLGSAVLL 612
D 1488 KTDITLATEVTKLPQTGDKTSFPRVILGGSANVLM 1520
```

RESULT 4

```
AB1347
probable peptidoglycan bound protein (LPXTG motif) lmo2178 (imported) - Listeria monocyt
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #ext_change 09-Jul-2004
C:Accession: AB1347
R:Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,
A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; WUID:21537279; PMID:11679669
A:Accession: AB1347
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1612 <GLA>
A:Cross-references: UNIPROT:Q8Y591; UNIPARC:UPI00000CFL1E2; GB:NC_003210; PIDN:CAD00256.1
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2178
```

```
Query Match 8.3%; Score 267; DB 2; Length 1612;
Best Local Similarity 21.5%; Pred. No. 1.1e-06;
Matches 145; Conservative 88; Mismatches 226; Indels 214; Gaps 26;

QY 35 NGEASQVVIHKKKMDLP-----PLIONSGKSEMF-----DKYGLADVTFSITNV 82
D 888 NNIISADRPINPTITDTMKGTGGAQVNVSSFKVMNETTGEDIDSKTY---DITTYDNPF 944
QY 83 TNEFYQRAAG-----ASVD-----AAKQAVOSLTPGKRVNO 114
D 945 TIQFDYKATAPIKYTYSTISLSGLVSNATTAASDYSGLPMTYKSRSTTSISPATFIS 1004
QY 115 GTTDAN-GNVTV-QLPKQNGKD-----AVYTIKEPREGVVAATN---MVAAPV--- 160
D 1005 GSGTATIGSLRITKVKDKNSKLTGAKFQLYTL-EGDKAGQEAATDSDGKIYMDGLQSG 1063
QY 161 -YEMIK-----QTDGSKYKGTBEELAVHIYKNNVAN---DGLSHYKVGNAHEGNA 209
D 1064 KTKLVETAPGTGTTISDEKDKGKEITVTADIASTVIENTEQGSVLLKEBSATYDATA 1123
QY 210 GAEFVSKSEG-----SPGTAVKIQ----- 229
```

```
D 1124 GAFFELQNDGTFVADNLVSNADGKLEVTDLARGDYQFVETKAPTVGLDGAFTKTFVE 1183
QY 230 -----GVKGLYTTWTDKEQAKRFITGKSYEIGENDPTAE-----NGGELT 272
D 1184 NOEAAVIVTKENTAKGSVVLTKEDSVSKATLSGAFFELQNDAGTGVKNTLTNAGKLE 1243
QY 273 VKNLEVSYLEEVRKAPNNAELIENOTKPTFEANNOPTVEKTVKNDTSKVDKTPPSLD 332
D 1244 VTDLAPGDYKFVETTKAPTVGEL--DAPVTFIIEFQATVAVNTKENVATGSVVLTKLD 1301
QY 333 GKDVAIGEKIKYQISVNIPLGIADKEGDANKYKFNLYDKHDAALTFDVTSGEYVALY 392
D 1302 SKRSNLAGAFELQKRGVLSLKD-----VTEANGQLIDN----- 1339
QY 393 DGDVIAIPENYQV--TEQANGFTVAANPAPIPLTGTGGLKRYFMHLNLEKADPTGPKN 450
D 1340 ----LAPGDYQVLEKAPTVGDLATPYEF-----TIEF-----NOKA----- 1373
QY 451 EAVNDGHTDDQTPPTVEVVTGSKRFIKVDGVTATQALAGASPVVRDQNSDTANYLKID 510
D 1374 -----PIQVTKTNSTGSGVLTGKTGEKA--LLANATFKLVDEDNVTENL--- 1419
QY 511 ETKAATWVKTAEMTPTTTADGLVDITGLKRYGYLEBTVAPDDYVLLT---NRIF 566
D 1420 -----TTDASGKLEINTLAPGQVLIETKAPAGYELDTVPVDVKITF 1461
QY 567 VNEOSYGTENLV 579
D 1462 DQKERTLQVTKTNL 1474
```

RESULT 5

```
AE1717
probable cell surface protein (LPXTG motif) [imported] - Listeria innocua (strain Clijf
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #ext_change 09-Jul-2004
C:Accession: AE1717
R:Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; M
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,
A.; Title: Comparative genomics of Listeria species.
A:Reference number: AE1717
A:Accession: AE1717
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1622 <GLA>
A:Cross-references: UNIPROT:Q92974; UNIPARC:UPI00000CC7F5; GB:AL592022; PIDN:CAC97509.1
A:Experimental source: strain Clijf1262
C:Genetics:
A:Gene: lin2821
```

```
Query Match 7.8%; Score 250.5; DB 2; Length 1622;
Best Local Similarity 20.7%; Pred. No. 9.4e-06;
Matches 171; Conservative 99; Mismatches 252; Indels 305; Gaps 32;

QY 14 LLLILPLFTSVLGTTFAPAEENGSAQVIHKKM--TDLPPPL-----IQNSGKMS- 64
D 884 LNIITPRLTSLPVITNRLQGNNTGAQPIKNSQVINTNBEINENENTDITREGNTFTI 943
QY 65 EFDKIQGLADV-----TFSITN--VTNE-----FYQRAAGAS----- 95
D 944 OFKNYTAAPIRKYNVSTISLSGPISNETTVEAEPSNVPMPKKNNVAVSPVFTVGTS 1003
QY 96 -----VDAKQAVOSLTPGKRVNOGTTDANGVTVOLPKQNGK 134
D 1004 GIATITGITTVNVEDDTTKLKGAKFQLYTLIDGKSGGELTNSGELT--LDGIQSGK 1061
QY 135 DAV-----YTIKEPREGVVAATNNVAPVYEMIKQTDGSKYKGTBEELAVHIYK 186
D 135 DAV-----YTIKEPREGVVAATNNVAPVYEMIKQTDGSKYKGTBEELAVHIYK 186
```

Db 1062 YKLVEFKAPGEGNISDEYKEG-----KEITVSSSG-----BEILLT---IK 1099
QY 167 NVNADSLHVKVKGTAENEGNGAEFVVISKEGSGCTKYIQGVKGLYTWTTDKEQAK 246
Db 1100 NAM-KKGVNLTTRKQDSASHEVLADAPFELQNAFGSK-----LKEKLI---TTD----- 1142
QY 247 RPIITGKSEYIENGDFTLEANGTEGLTVKNLEVGSYLLEEVKAPNNALLENQTKPTPTIE 306
Db 1143 -----ASGNIEITDLAPGDYKLIETKAPTQGL--DAPVNFITID 1180
QY 307 ANNOPTVEKTVKNQTSKVDKTTPSLDGKQVAIGEKIKQISVNIPLGIADKE--GDANKY 364
Db 1181 F-NQSEAAKVSKTNTAKTG--TVMELTKQDSA-----TNALADATFELRNEDETL 1227
QY 365 VKENLVDKHDAALTVDNVTSGEYAAVALYGDVTVIAPENYQVTEQANGFVAANPAYIPTL 424
Db 1228 VARENLTVDNGEISVADLAPGDYKLI-----ETKAPTQGLDAPVNFITIDNOSVEKV 1282
QY 425 TPGGLTKFVYFMHLN-----EKADPTKGFKNKANV---DNQHTDDQTPPTVEVVTGK 474
Db 1283 SKTNNAKGTAVLTTRKQDSATNLEADATFELNESPDLVRENVLVTDNGEISVADLAPG- 1341
QY 475 RFIKVDGDTATQALAG-----ASFVVRDQNSDTANTYKIDETTKAATWYKTKAEATT 527
Db 1342 -----DYKLIETKAPAGYQOLDATPVNFITIDFNOSEAAKVSIN-TAKTGTVLTTRKQDSAT 1395
QY 528 FTTTAD-----GLVDITGKXGTYLLEETVAPDDVLLTNRIE 565
Db 1396 NAELEADATFELNESPDLVRENVLVTDNGEISVADLAPGDYKLIETKAPTQGLDAPV 1455
QY 566 FVV-----NEQSYGT----- 575
Db 1456 FTIDFNOSEAAKVTKTKKKKIGTIIIVNFIDVDGNQNDKSVHTGNVGEYSVKAKEIVGH 1515
QY 576 -----TEN 578
Db 1516 TLVKDPANKKMYKETSQETITVEYKEMKAPITVEPTPEPSKTEOLTESATVAEPKIKON 1575
QY 579 LVSPEKVPKHKHGTLPSTGKGIVYVIGSGAVLLLAGVYFAPRRKE 625
Db 1576 IKTTKSPN-NRKKLPSITGDEPPYTMFTG-LFVSVAAGVFLRKPKQ 1620

RESULT 6
AF1717
probable peptidoglycan bound protein (LPXTG motif) lin2282 [imported] - *Listeria innocua*
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AF1717
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Entlian, K.D.; Fehrl, H.; Jones, L.M.; Karet, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Meak, C.; Schlueter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative Genomics of *Listeria* Species.
A:Reference number: AB1077; MUID:21537279; PMID:11676669
A:Accession: AF1717
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1806 <GLA>
A:Cross-references: UNIPROT:Q929U3; UNIPARC:UPI00000CC7P6; GB:AL592022; PIDN:CAC97510.1;
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin2282

Query Match 7.7%; Score 247; DB 2; Length 1806;
Best Local Similarity 22.0%; Pred. No. 1.7e-05;
Matches 135; Conservative 63; Mismatches 219; Indels 198; Gaps 20;

QY 27 TTTAPAEENGSAQVLIHKKKMTDLBPDLIONGSKGMS-EPDKYGLADVTTSIVYVME 85
Db 1177 TPTKTFVEBQTAIVTK-----ENTAKAGSVVLTQDSATATATIA---GAE 1221

QY 86 FYEORAAAGSY-----DA-AKQAVSLTPGKPVAGQTTDANGNTVOLPKK-----ONG 133
Db 1222 PELQKADGTKISENLVTTBADKIEVNDLAPGQYQVETKAAAGVLDSPSPFTIEFNQD 1281
QY 134 KQAVYTIKEEPKEGV-----AATNVVAFPVV-----MIKQTDG----- 169
Db 1282 KAAVTKENTAKTGSVVLTKQDSVTKATLAGEFELQKADGTKSENLVTTVDGIEVND 1341
QY 170 -----SYKKGTELAHVHLYPKKVVANDSGLHVKKVGTLEN 205
Db 1342 LAPGQYQVETKAAAGVLDTPYKFKVFNQTSVAVAKEN-TAKAGSVLTKQDSVTK 1400
QY 206 EGLNGAEFVVISKEGSGPGTVKTYIQGVKGLYTWTTDKEQAKRFTIGKSEYIENDFTEAE 265
Db 1401 AALSAPFELQTAITTK-----VKONLT----- 1424
QY 266 NGTGLTVKNLEVGSYLLEEVKAPNNALLENQTKPTPTIEANNOPTVEKTVKNDTSKVD 325
Db 1425 NASGEIEVADLAPGDYKKEVETKAPTGYEL--DAPVTFTEFNQATAVKVTKEVNAKTGS 1482
QY 326 KTTPSLDGKDAVIGKIKQISVNIPLGIADKEGDANKVYKFNLYDKHDAALTFPDNVTSG 385
Db 1483 VVLTKLDSKSKSNLAGAEFELQTKLGVSLDKD-----VTEANGQLOINLAPG 1531
QY 386 EYAAVALYGDVTVIAPENYQVTEQANGFVAANPAYIPTLTPGGLTKFVYFMHLNEKADPT 445
Db 1532 DYQLV-----ETKAPTGYELDAPVETIEFN-----QKQAV 1563
QY 446 KQFKNBANVDNGHTDDQTPPTVEYVYTGKRFKVDGDTATQALAGASFPVRDQNSDTAN 505
Db 1564 -----QVTKTKMSTGSVVLTKTGQTKA--PLADATFKLVADANNVTE 1605
QY 506 YLKIDETTKAATWYKTKABATFTTTADGLVDITGLKXGTYLLEETVAPDDVLLTNRIE 565
Db 1606 NL-----TTDASGKLEITNLAPGDYKLIETKAPKGYELDTVPVD 1644
QY 566 FVVNEQSYGTENTLV 580
Db 1645 VKI-----GFQNOQV 1654

RESULT 7
A32347
fimbrial protein type 2 precursor - *Actinomyces naeslundii*
C:Species: *Actinomyces naeslundii*
C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 05-Oct-2004
C:Accession: A32347
R:Yeung, M.K.; Clisar, J.O.
J. Bacteriol. 170, 3803-3809, 1988
A:Title: Cloning and nucleotide sequence of a gene for *Actinomyces naeslundii* wv045 typc
A:Reference number: A32347; MUID:88314866; PMID:2900829
A:Accession: A32347
A:Molecule type: DNA
A:Residues: 1-534 <YEU>
A:Cross-references: UNIPROT:P12616; UNIPARC:UPI000012MA55; GB:M21976; NID:q141847; PIDN:
C:Superfamily: fimbrial subunit precursor
C:Keywords: transmembrane protein
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-534/Product: fimbrial protein type 2 #status predicted <MAT>

Query Match 7.6%; Score 244.5; DB 2; Length 534;
Best Local Similarity 23.4%; Pred. No. 3.9e-06;
Matches 147; Conservative 63; Mismatches 220; Indels 197; Gaps 29;

QY 90 RAAAGSVDAAKQAVSLTPGKPVAG-----TTDANGNTVOLPKKQNKDAVY--- 138
Db 10 RRAAAAAGVLTAAVIGLA---FMAQAEANMHDINTALGSLTIH--KRLNDGNPIGAP 64
QY 139 --TIKEEPKEGVAAATNVVAFPVVEM-----IKQDGSYKGTETELAVVHIYKVVVAND 192
Db 65 DGTASNDGKAPVSG--VQFTAYEINGIDLKTSBGNAK--VNLVTNGALPDAACNP 119

```

Oy 193 GSLHKKKGTJENBENAGAEFVYSKESBSPGVKXIGVMDGLTWTWTDKQARPFITK 252
Db 120 G-----QPTLPVYTFRRSRVSGD-----TDGR-----141
Oy 253 SYEIGENDPTEAENGOTGELTVKNLEVGSIYLEEVAKPNNALTEHOTKPTIEANNOTP 312
Db 142 -----GEAKIESLPVKAYLVCETKTPGNIV-----QAKAPVVVITPHNT 181
Oy 313 VEKT-----VKNDTSKVDKTPTPSLDGKVALGEKIKQISVNIPLGIADKEGA 361
Db 182 AAKKDGTWLYOVHVYPRKIEKIEVAKTIEDQNNNGVYISKVRFPFSSLP-----KLD 235
Oy 362 NKTYKE---NLVDKHDALTPDNVTSG-----EYAYALYDGYIAPENQVTEOA 409
Db 236 NSYYKXYQPKOTLDNRKQVTAIVDTVLGGTRLDEGTDYLTGT-DGQTV-----TVTFNQ 288
Oy 410 NGFTYAVNPAIPIPTLRGTLKFVYFMHINKAD-----PTKGFNKE 451
Db 289 NGLSLDKG-----NPGOKLOAVEGVSEVGOSINNTAOLISDPTYAQBPAPEPP 341
Oy 452 ANVDNGHTDDQTPTEVEVVT---GGKRFIKVDGVTA--TOALAGASFV---RDONSDT 503
Db 342 ANPDN-----PPTTEQVTSKMGDLTIKKVGDGDRSDKXGGLKAEPQIKAKADAYDT 394
Oy 504 ANYLKIDETTKAATWVKTKAEATFTTTADGLVDITGL-----KXGTY 546
Db 395 C-----SPBADQPLTINGESTFTTGEGTINFKALLFVSDVODTGRDNRVDAPRXY 447
Oy 547 YLEETVAPDDYVLLTNRIEFVNVNOSYGTENLV---SPEKYPNKHKTLPSTGGKGIYV 603
Db 448 VLVTETKAPAGVYLPADADASRAITVEPGAGVTOQVIVDNVKQSPG-----LPLTGANGMLI 502
Oy 604 YLGSAGVYLLIA--GVYFA---RRRKEN 626
Db 503 LTASGAALMTAVGSVLVARYRRKRPN 529

```

```

RESULT 8
AH1094
probable peptidoglycan bound protein (LPRNG motif) Imo0159 [imported] - Listeria monocytogenes C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence revision 27-Nov-2001 #text change 09-Jul-2004
C/Acession: AH1094
R/Glaser, P.; Fraigneul, L.; Buchriesser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloekerke, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fahni, H.
D.: Jones, L.M.; Katzev, U.
Science 294, 849-852, 2001
A:/Authors: Kretz, J.; Kulm, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schlueter, T.; Smoes, N.; Tierce, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.
Article: Comparative genomics of Listeria species.
I./Reference number: AB1077; MUID:21537279; PMID:11679669
A/Acession: AH1094
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-793 <GLU>
A/Cross-references: UNIPROT:O8YAG7; UNIPARC:UPI0000054CA9; GB:NC_003210; PIDN:CAC98374.1.1
A/Experimental source: strain BGD-e
A/Genetics:
A/Gene: Imo0159

Query Match          7.5% Score 242.5 DB 2 Length 793;
Best Local Similarity 20.9%; Pred. No. 9.1e-06;
Matches 177; Conservative 96; Mismatches 286; Indels 287; Gaps 32;

    Oy      4 LKRYVTVSTLLILPLF--TSVLGTTAFENGESAAVIHKKKMTDPDLIQNSGK 61
              ||| |::||::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db       8 LKYGLAPLSILIVASTIFQITTVKPATSYSGSOPLATVE-----LIDNDGV 53

Oy      62 EMSEFDKROGLADV--FPSIVNTN-----EFYQRAAGASV-----DAKQAPOS 105
           ::||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     54 POTDRGYIDNM-DHAYTISINSTNVKAAGDTMDPFLPSQLALADPADVDKSIGQVVGT 112

Oy     106 LTPEKPAVAQ-----GTTD-----ANGNVTVQLP-KKONG-----   133
```

```

Db 113 ATVKAQNOUUTLVSDYVEKSHDVKGELDLWTAENQKVIIGNETVLDLVPLENGTIVIDV 172
Qy 134 ----KDAVYTIKEEPKEGVVAATNMVAVPV-----YEM 163
Db 173 EVGEKTPVSPTELEFKYGWDASNPSLIHVVVRVNAKVNI PNAVFTDIIGAKOTLNFDS 232
Qy 164 IKQTDGSKYKGTBEIAVNIYKKNVA-NG-----SLHVKKVGAENEG----- 207
Db 233 IKAHGHTYSADRVFTAAPISSSTWFSATSDGFSVALGNLTDSVOJSTTTTTOGASTOY 292
Qy 208 ----LNGAEFVLSKSE-----GSPGVKVIYIOGVKDLVYTTWTDKBOAKFTIGSKY 254
Db 293 DNPAKLAGTDFVTKQJSTWTPPASGGGEG-----GGTGSVTLTKEDATKATLBGAFF 346
Qy 255 EIGENDTTEAB-----NGTGBLVKOLEVOSYILBEYKAPMNAELIENQKTPPTTEANN 309
Db 347 KLVDSKSTVLOENITTTNASSGOLSIADLKPTYOILFKAPGYKL--DTTPVEFTIGENN 404
Qy 310 QT-FVEKTVANDTSKVDKTPPSLDGKOVALGEKIKYOISVINIPLAGIADKEGDANKVKFN 368
Db 405 QATVYTKENTLNTGSEVELTKLDAATKATLAGATFELD-----DKEGNT---LQTD 451
Qy 369 LVYDKHAAITFDVNTSGEYAVALYDGT-----VIAENYQVTEQANGFTVAV 416
Db 452 LKTDENGVLKATDLVPSGYQFVETSAPGYKLDNSPVSEFVIAGETDQVVKTKENTLEV 511
Qy 417 NPATIPPLTUG--GTLKFPVYFMHLINEKADPT-----KGFRN-----EAN 453
Db 512 GSVELTKLDSATKATLAGATF-ELQDKEGNTLQTDLKTDEGVLKATDLVPSGYQFVETS 570
Qy 454 VDNGHSTDDQTPPTVEAVTG--GKRPIKYDGDVT-----ATQA-LAGASFFVARD 498
Db 571 APGYKLDNSPVSEVAVAGETDQVVKTKENTLEVGSEVELTKLDSATKATLAGATFELD 630
Qy 499 QNSDTANYLKIDETTKATVWTKAETTTTADGLVDTIGLKGTYIYIEYVAPBDYV 558
Db 631 KEGNTLO-----TGLTDEGVLKATDLVPGYQFVEKTAPIGYE 670
Qy 559 LLTNRIRFVNNEQSGYTTENLV-----580
Db 671 LDTTPVPSF---EIVAGETDPIVAKTKENTLVPPRPVPTPVPRPVPPTVPPTPPLBPVP 727
Qy 581 -----SPEKVENKHKGTLPSTGKGIIYVYIGSGAVLLIIAG 616
Db 728 YEFTVPTKBEVVPVTKTENSDESPKTEPIRITQSLPKTGDTNPSFAGLIVILALISLG 787
Qy 617 VYFARR 622
Db 788 LLLKRX 793

```

RESULT 9
 F83901
 hypothetical protein BH2014 [imported] - *Bacillus halodurans* (strain C-125)
 C|Species: *Bacillus halodurans*
 C|Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C|Accession: F83901
 R|Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A|Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A|Reference number: A83650; MUID:20512582; PMID:11058132
 A|Accession: F83901
 A|Status: preliminary
 A|Status: preliminary
 A|Molecule type: DNA
 A|Residues: 1-1816 <STO>
 A|Cross-references: UNIPROT:Q9KB94; UNIPARC:UP100000C3D8A; GB:AF001514; GB:BA000004; NIT
 A|Experimental source: strain C-125
 C|Genetics:
 A|Gene: BH2014

Query Match	7.3%	Score 234.5	DB 2	Length 1816
Best Local Similarity	20.2%	Pred. No. 8.5e-05		

Matches 126; Conservative 80; Mismatches 221; Indels 197; Gaps 21;

```
Qy 40 QLVIRHKKMTDLPPELLIONSGKMSBPDKYQGLADVTFISYVNTNFEYFORAGASVDAA 99
Db 960 QSTVNTAKIVDPSP-----NQILBEDSFQ-----LFTVNTATN----- 993
Qy 100 KQAVOSLTPGKPVNAQT-----TDANGNTVTVQLPKQNGKDAVYTIKEEPKGVVAAT 152
Db 994 ---GTTKGAPLERGVADYILEIFTEDEGNETFEL-SPPNVIDRAYTL--EYKSFYIAQN 1046
Qy 153 NMVVAFPV-YEMIKOTDGSYKGTGTELAIVHVIYKPVAVND-GSLHVKVGTAEINEGANG 210
Db 1047 NEFVSNQVSPFGDELTMTNESTNSSISQVRLSGSGTSGRGSGISLVTKVDATSGETIAG 1106
Qy 211 AEPVISESGSPGVVYIKQGVKGLYVTT--DKEQAKRPITGKSVEIGEND----- 260
Db 1107 ATPVLVYSTGTEIVLRTVVTG-EDGIATPANIMYDYLIRETSAPEGYLVGINDESHITLN 1165
Qy 261 -----FTEAENGTE-----L 271
Db 1166 TELHELVENAKIIRDVQLHKIIDETGSLQCAVFTLQORQSDPFVDYRENLETFDENGV 1225
Qy 272 TVKNLEVGSYILIEBKVAENNAELIENQTKTEFTIEANNQTEVEKIVKN-----DTSKV 324
Db 1226 AVEDLEPDQYQVEVAAPSGVYL--NADPIFSIGEQTGTEVILTKENEIILGSEVELKV 1283
Qy 325 DKTPPS--LDGKDAVIGEKIKYQISVNIPLGIADKEGDAKRYKFNLYDKDAALTPDNV 382
Db 1284 DRDSSSYLLEGAE-----FSLDEGDNV--LQGLTTDEGKLVVSDL 1324
Qy 383 TSGEAYVALYDGDVIAPENYQVTEQANGFVAVNPVAPITLTPCGTLKFPVFMHLEKA 442
Db 1325 RPNYQPV-----ETTSPEHLDDTPYRPFITVLDDQQLVYT----- 1362
Qy 443 DPTKGFKNEAVNDGHTDDQTPPEVENVYTGGRKFIKVDGVTATQALAGSFVVRDQNSD 502
Db 1363 -----ATNRLIPGSVQLKVDADHPAV-VLEGAFESFLDEDCN 1399
Qy 503 TANYIKIDETTKAATWVTKAEATPFTTAGGVINDITLXKGYLYEETVA PDDVYVLTN 562
Db 1400 IVQ-----EGLSTDANGQLIVTDLRPNYQVPEVETAPADEYILDAT 1439
Qy 563 RIEFVNEQSYGTENTLVSPKVP 586
Db 1440 PREFTIERSQEGISIVENELIP 1463
```

RESULT 10

AC1458

surface anchored protein (LPXTG motif) [imported] - *Listeria innocua* (strain Clj11262)

C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C:Accession: AC1458

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, A.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.; Jones, L.M.; Karet, U.

D: Science 294, 849-852, 2001

A:Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapatk, G.; Madeno, E.; Maitournam, A.; Meakins, C.; Schluecker, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, O.; K. C.; Schluecker, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, O.

A:Title: Comparative genomics of *Listeria species*.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1458

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-681 <GLA>

A:Cross-references: UNIPROT:Q92PA4; UNIPARC:UPI00000CC163; GB:AL592022; PIDN:CA95435.1

A:Experimental source: strain Clj11262

C:Genetics:

A:Gene: lin0202

QY 4 LKKVWT-VSTLLILPLF-TSVLGTTTAAEENGESAOVLHKKQTDLDPDILLQNSGK 61
|||:::|:
Db 8 LKRVGIAFLPSLVLIASIVPOTTIVKATSYGSEPLNVE-----LIDKDT 53
|||:::|:
QY 62 EMSEFDKYGOL-ADVFSIYNVTNEFEQRAAGSVDAAKOVSLTPEGKPVAGQTTDAN 120
:|||:::|:
Db 54 PKTDREYIDNMNVHYTWISIPNSTNV---KAGSMDFALPS--QLALATDLAFAVKSK 106
:|||:::|:
QY 121 GNV--TVQLPKONGDAVYT-----IKEEP-----KEGVAAITMV-VAPPYEM 163
|||:::|:
Db 107 GGTGATATKVRATNQUTLVFSDYEVEKSHDIKGELDFWTFENOKYITGENEKINLEPIENS 166
|||:::|:
QY 164 -----IKOTDSYKYGTGE-----LAHYIYPNNV----- 188
|||:::|:
Db 167 TINVDNVGEKTPVSPEETIFKTGWANDANNPSLIHWVVRYVAQKNIFNANFTDIIGAKQ 226
|||:::|:
QY 189 -----VAN-----DG-----SLHYEKVGAENE 206
|||:::|:
Db 227 TLNPDSIKAFHGYSYTRDVFTAGPTISNTNFASKTSDDGPSVTLGNLTDSVOJSYTTATD 286
|||:::|:
QY 207 G-----LNGAEFVISKSE-----GSPEGVIKIQQVKDLTYTTTDEQAK 246
|||:::|:
Db 287 GKSTGYDNRTAKLSGDEFVTKQTSTWTPASGGGGADGTT-----GSTVLTUKODAKTK 338
|||:::|:
QY 247 RFTTGKSYE-----IGENDFTEANGTGLTIVKNILEVGSYLIEEVKAAPNAELIENO 298
|||:::|:
Db 339 ATLEGAEFLVDANGTILOGENITTD-----SGQLANSMLKPATYQULIERKAPTGYTL--DS 393
|||:::|:
QY 299 IYTPPTI-EANNOTPEKVTKNDTSKYDKTTPSLDGKDAIGEIKIQOISVINPIGIADK 357
|||:::|:
Db 394 TEVEFTIGENNKEITYTKENTLDTGAVELTKVDATAKAILAG----- 435
|||:::|:
QY 358 EEDANKYKVENLDVKDADALTFPDNTSGEYAYVALYDQDTVAPENVQYTEANGFTVAVN 417
|||:::|:
Db 436 -----ATFELDCKGSIYLODKTIDENGILAKTD---LVPGNYQ----- 471
|||:::|:
QY 418 PAYIPTLRPGTTLKFVYFMHLNEKADPTKGFKN EANVNGHTDOTPEPTEVV--TGGR 475
|||:::|:
Db 472 --FVETSAPFG-----YKLDSPPVFELIAD-----TDQIVKTKENILEVGA 515
|||:::|:
QY 476 FIKVDGDVARTQA-LAGASFVRVDQNSDTANYLIKIDETTKAATWKTAEAFTTTTAG 534
|||:::|:
Db 516 LTRKD--AAFKALLAGATFELODKQKTLQ-----TGLTIDENG 552
|||:::|:
QY 535 LVDTIGLKGTYYLSEETVAPDDVYLLTNRIEFVNNEOSYGTENLVSEKVPNKHKGLTP 594
|||:::|:
Db 553 VLKQTDLVPGTGVFVETNAPIGYELDSPVFE---EIYIGNNQVVKYTK-ENKAKETTP 608
|||:::|:
QY 595 ST 596
|||:::|:
Db 609 PT 610

RESULT 11

G86643

[hypothetical protein ybf (imported) - Lactococcus lactis subsp. lactis (strain IL1403)]

Cyspecies: Lactococcus lactis subsp. lactis

Cdate: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

CAccession: G86643

R.Bolotin, A.; Wincker, P.; Mueger, S.; Jalllon, O.; Malarme, K.; Weissenbach, J.; Ehrlich, B.

Genome Res. 11, 731-753, 2001

Affiliate: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s.s.

A.Reference number: A86625; KWID:213518e; PMID:11337471

A.Accession: G86643

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-1983 <STO>

A.Cross-references: UNIPROT:Q9CJ55; UNIPARC:UPI000000C676A; GB:AE005176; PID:g12723000; F

A.Experimental source: strain IL1403

A.Gene: ybf

Query Match 7.1%; Score 228; DB 2; Length 1983;
Best Local Similarity 21.9%; Pred. No. 0.00022;
Matches 113; Conservative 63; Mismatches 152; Indels 186; Gaps 22;

107 TFGKP---VAQGTDTANGN-----VTVOL---PKQNGKDAVYTIKEEPKEG---VVA 150
1100 TGKDKDKLIREGDTDKNGDLKLGKGLKLVTDYLLIETKAPG---YTLDEELVYGGKTIYA 1155
151 ATNNVAVPVEVEMIKOTDSYKKGTBELAVNHYPKNVAVNDGSLHVKVYG--TAEN--- 205
1156 KDGVSNNPFIQVENE-----PKVI-----LKKVGLTIENGKEV 1190
206 -GLNGAEFVISKSESGPCTVKYIQGVKGLYTTTDEKQAKRFTGSKSYEIGENDPTA 264
1191 KNPLOGAEFKVLDSNQ-----VSGYEKLTSD----- 1218
265 ENGTEGLYVKNLEVGSYILEEVKAPNNABLIENQYTPPTTEANNQTPVEKTVKNDTSKV 324
1219 --SSGNVTIEKLPKGSIVETKAP--AGYILDPTPIDELKNEBGII----- 1263
325 DKTTBSLDGKVAIGKIKYQISVNI-----PLGIADKESDANKVKNVLD 371
1264 -----PDINL--EKVYQSAQLIKNSNGQALSGAIFKVIDKQNT--IQTNLS 1310
372 KHDAAITPDNWTSGEYAVLYDGDVIAPENYQVTEQANGFTVAVNPAYIPTLTGGLTK 431
1311 DDDGKVTATGSLSPGDSFY-----ETKAPGTGILNTIPHTISA----- 1350
432 FVYFNLNEKADPTKPKNEANVNDGHTDQTPPEVVTGGRFIKVDGVTATQALAG 491
1351 -----BEEGQPOVNIASDNFIN-----YQSAELIKQDNE--GGPLSG 1386

492 ASFVVRDQNSDPAVNIKIDETKATWTKAETFTTTAGLVDITGLKGTIYLET 551
1397 AEFKIVD-----KAKTIQA-----NLTEGKQKVVVDLAPGEYSFVET 1426
552 VAPDDVLLTNRIEFVNVNOSYGTENLVSPKVPN 587
1427 KAPGTIILNTNVPVHTISDKERGQRTTIASDSFN 1462

RESULT 12
A35400
surface protein T6 precursor (strain D471) - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C>Date: 31-Aug-1990 #sequence_revision 31-Aug-1990 #text_change 09-Jul-2004
C:Accession: A35400
R:Schneewind, O.; Jones, K.F.; Fischetti, V.A.
J. Bacteriol. 172, 3310-3317, 1990
A>Title: Sequence and structural characteristics of the trypsin-resistant T6 surface pro
A:Reference number: A35400; MUID:90264329; PMID:2188957
A:Accession: A35400
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-537 <SCH-
A:Cross-references: UNIPROT:P18481; UNIPARC:UPI00001368BD; GB:M32978; NID:g153842; EIDN:
C:Keywords: transmembrane protein

Query Match 6.8%; Score 220.5; DB 2; Length 537;
Best Local Similarity 22.4%; Pred. No. 8.4e-05;
Matches 128; Conservative 82; Mismatches 185; Indels 177; Gaps 31;

155 VVAPVVEVEMIKOTG---SYKGTBELAVNHYPKNVAVNDGSLHVKVYGAENEGANG 210
40 VTLVYKIGGVNTNDSPFINFRY-ARGVSLTGTGPS-----QGITTIAN-GINT 87
211 AEFVSKSEG---SPGTVYI---QGVKGLYTTTDEKQAKRFTGSKSYEIGENDPTA 264
86 GKIKPFTENVSISNGIATYNNAGASVYIALTGATDERTVPIILASV----- 137
265 ENGTEGLYVKNLEVGSYILEEVKAPNNABLIENQYTPPTTEANNQTPVEK--TVKNDTSK 323
138 -NGEKNLVTKNIDSKSNYL-----YQGTSAKSLSPITK 172

324 VDKTTPSLDGKVAIGEKIKYQISVNIPLGIADKESDANK--YVKNELVDKHDALTFDN 381
173 VTGTDVVKKTTSLASVLSYSTLELP--SYTKEA-VNKTIVVSDNM--SEGLTNNFS 227
382 VT---SGEYAVLYDGDVIAPENYQV---TEQANGFTVAVNPAYIPTLTGGLKVFYE 435
228 LTVEMKQKNAINTBDSVNV--ENTKIGIAKEVNNQFNLSFYDLSLSISPNISYKAV-- 283
436 MEHLEKA-----DPTKGFKNBAND-----NGHTD----- 460
284 --VNNAIIVGEEGNPKAKEFFYSNNPTKG--NTYDNDLRKPDKNGKITSKEDSKIVTYOI 340
461 -----QTPPEVVTGGRFIVDDGVNATQALAGSFVVRQ 499
341 AFRKVDYSKTPDIGAIFGVYDTSNKLIDIVTNK-----NGYASTQVSSG-KYKIKEL 394
500 NSDPAVNIKIDETTKAATW---VTKA--EATTEPT---TADGLVDITGKYGTV--- 546
395 KAPKGSINTETFEITANNVTATVTSANSKSTYTSQKNAKDNEQGVMLKNGIFYSI 454
547 -----YLBETVAPDDVLLTNRIEFVNVNOSYGTENLVSPKVPNKKGLTP 594
455 DSRPTGNDVKEAYIISTKA-----LTDGTFESKSESGT---VLEETDIPNKLGEIP 505
595 SNGKGIYVY--LGSQAVLLIAGVYFARRK 624
506 STGSIQTYLFAIGSAMIGAI-GIYIVRRK 536

RESULT 13
AB3695
hypothetical protein BH0361 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: AB3695
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sakaki, R.; Maeni, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: AB3650; MUID:20512582; PMID:11058132
A:Accession: AB3695
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1661 <STO>
A:Cross-references: UNIPROT:Q9KFW0; UNIPARC:UPI00000C385A; GB:AP001508; GB:BA000004; NIT
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0361

Query Match 6.7%; Score 216.5; DB 2; Length 1661;
Best Local Similarity 20.2%; Pred. No. 0.00074;
Matches 140; Conservative 74; Mismatches 225; Indels 253; Gaps 25;

62 EMSRFDKYQG--LAQVTSIVYVNEF---YEQRAAGASV----- 96
1088 EYTKVDADQGVYLGATFTLVDSBGEFARTLETGEDGATFVNLGYDLKEDSAPBG 1147
97 -----DAKQAVQSLTPEKPAQGTDTAN-----GNTVOLPKQNGK----- 134
1148 YLVGINDTORVYIDVTLHEVTYENKSDINRVSANGAVQLQRYVDEBTGSLQALPALQ 1207
135 ---DAVYIT--KEEPKEGVNATNNVAVPVEVEMIKOTDSYKKGTBELAVY----- 181
1208 KYDDEFTVIAEMETBEGIVFAGSLB--PGDYQFVELNAPVGYKLDFTPVVFTVEEDRT 1264
182 -----HYPKNVAVNDGSLHVKVYGAENEG--LNGAEFVISKSESGPCTVKYIQGV 231
1265 ETEBLQKENVHLP-----GSVQLVKVDADDAATLBSAEFTLLDGBGVN-----V 1309
232 KQGLYTTTDEKQAKRFTGSKSYEIGENDPTAENGTEGLYVKNLEVGSYILEEVKAPNN 291
1310 OEGT---TTDE-----NGQVVVTDLKPGEYQFVETKAPAG 1341

```

QY      292 AELLNQTKEPFTLEANNQTPVEKTVKNDKSRVDKTPSLDGNVALGEKIKYQISVINP 351
Db      1342 YEL--EAPRIGFTTIERNOQEVATVAVENHLLPGSVOLVKVADDAAILLE-----GAE 1392
QY      352 LGIADKREGDANKYKFKFNLDKHAALTFPDNVTSGEYAAVLXDGPTVLAPENYQVTEQANG 411
Db      1393 FSLLDGEGNV---VQEGLTTPATGQVVVTLDKPEFYQV-----ETKAPAGYEIDAPIG 1444
QY      412 FTVAVNPAYIPFTLTPGGTLKFFVFMHLEKADPTKGPFKEANVDNGHTDDQTPPEVEVT 471
Db      1445 FTIERNOQEVATVA-----VEN-----HLIP 1465
QY      472 GGGKRFIKDGSVNTAQTALAGASFVVRDQNSDPTANYLKIDETTKAATVYKTAETATPTTT 531
Db      1466 GSVOLIKDKRDSAV-TLEGAFSLDDE--GNVLR-----EGIRTG 1504
QY      532 ADGVLDTGLKKGYYLEETVAPDDYVLLTNRIEFV-----568
Db      1505 ENGQLMVLDLRPGDYQFIETKAPFGYELDDTPPIAFTIEKGTDEVLTLVFNKLTPEVPES 1564
QY      569 -----NEQSYGTENLVSP-----KVPNKHKG-----591
Db      1565 PENPENPENPENPEIPEDLETPESPENPETPESPKHSDESVFNSDKNPEKSSDNNRID 1624
QY      592 TLPSTGSGKITYVYLGSGAVLLLIAGYPARKR 623
Db      1625 RLPTQGEFFLFTLLIGLLIMTAGVLLLKKR 1656

```

RESULT 14

hypothetical protein yhgG [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revistion 23-Mar-2001 #text_revision 09-Jul-2004
C:Accession: F86779
R:Botolint, A., Wincker, P., Mauger, S., Taillon, O., Malarme, K., Weissenbach, J., Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: AB6625; MUID:21235186; PMID:11337471
A:Accession: F86719
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-614 <STD>
A:Cross-references: UNIPROT:O9CHH3; UNIPARC:UPI00000D43FC; GB:AB005176; PID:g12723675; E
A:Experimental source: strain IL1403
C:Genetics
A:Gene: yhgG

Query Match	6.7%	Score 215;	DB 2;	Length 614;
Best Local Similarity	21.3%	Pred. No. 0.00021;		
Matches 155;	Conservative 79;	Mismatches 269;	Indels 226;	Gaps 38;

Qy	1	KKOLKKVYVYVYVSTLLILPLFTSVLGTTAAE-----ENGSAQLVHHKKMTDLP	52
			:
Db	3	LSNLSKNPALMSVLSLLALSTLAGFSLATVNAKTTNVADQVNSGEVA---IYAQSAAGNQ	59
Qy	53	DP-----LIONSCKEMSE-----FDKYQGLADVTF-----SIYVNTNE	85
		:::	::
Db	60	NPTQNGGLVDSQGNFVNSGIANNGTGGQAGNTSGTKNDGSOQTLLBGSFSTPMANVT--	117
Qy	86	FYEQRAAGASVDAAKQAVQSLTPGKPVYAQCTTDANGANTVOLPKKQNGKDAVYTIKKEEP	145
Db	118	FSATKRYVTGAVPTG-VTDPSFTDITSVAPVTTDASG-----LADFTGLTDYVLFHQ---	168
Qy	146	EGVVAALTMNVVAFPYFEMIKQTDGSYKKGTEBLAVYHLYPKNVVANDSLHVKKVGTLEN	205
		:::	:::
Db	169	---VTVVNGITT--VGDFIVQV---SHBDSQAGIVNVYPKLDMSSSAGLSTSATTNND-	218
Qy	206	EGLNG-----AEFVLSKSGSPQTVKYIGVVDGAYTWT--TDKQAKRFI	249
		:::	::
Db	219	DNFNGQOTNOJLANPNATGNSDQTLTNTDNNANENRLANG-----TWTNGSDNQVTTTAA	272
Qy	250	TGKSYKEIGEN--DPTBEANGTEGELTVKNLLEGVGSYLEEVAPANNAAELIENQTKPTIE	306

[illegible]

RESULT 15

A:UniProt: A1744
 internalin protein, probable peptidoglycan bound protein (LPXTG motif) homolog lin2495
 C:Species: Listeria innocua
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AB1744
 R:Glasner, P.; Prangnell, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Fehli, H. D.; Jones, L.M.; Karet, U. Science 294, 849-852, 2001
 A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurupkat, G.; Madueno, E.; Maltournam, A.; Mok, C.; Schlueper, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A. Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID: 21537279; PMID: 11679669
 A:Accession: AB1744
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-940 <GLA>
 A:Cross-references: UNIPROT:Q928N7; UNIPARC:UP100000CC6B3; GB:AL592022; PTDN:CA937722.1
 A:Experimental source: Strain Clp11262
 C:Genetics:
 A:Gene: lin2495

```

Query Match      6 6%; Score 212.5; DB 2; Length 940;
Best Local Similarity 23.5%; Pred No. 0.0053;
Matches 160; Conservative 77; Mismatches 255; Indels 189; Gaps 37;

Oy 18 LPLFTSVLGTTFAAFEENGESAQLYIHKKQMTDLDPDLPIONSGKMSFEFDKYQGLADYTF 77
Db 370 VPLSYSIKXQPIKXSEKTNASAVNFYKXDNDELAPSETTSGK--SGEYIQ--TTEKTI 425%
Oy 78 SIYVNTVEFEQCPAAG--ASVPAAKOAVOSLTPGKPVQQTITDANGVTVQLPKKONGK- 134
Db 426 TNYKAKR--IEGASQGFSDPTTYYVYERKADGAPVTVKTVYDQGN--ELATSDTLNGKI 482%
Oy 135 DAYV-----TIKEEP--KEGVNATNNVVAFPYEMIKQTDGSKYKGTBELAVH 182%
Db 483 DAPVOSTASITGMAYKTTTPAANGVFTNANQTVYV-YERKADGAPVAVKTVYED----- 53%
Oy 183 IYKGVNVVANDGSLHKV-----KVGTAENBGL-----NGAEFYISSEG 22%

```

```
Db 537 ---GNELATSDTLNGKIDAPYOSTAKSITGMWAVKTPPANANGVFTNANQTVTVYVEKADG 593
Qy 221 SPGTWKYIGVNDGLYTTWTDKEQAK-----RFTGKSIEIGENDPTEAENGTEL 271
Db 594 APVTWKYVD--BDGNELATSDTLNGKIDAPYOSTAKSITGMWTVK-----TPPANANGVF 645
Qy 272 TVKNLEVGSYILEEVK-APNNAELIR---NOTKPTIIBANNQTPVEKTVKNDTSKVDKT 327
Db 646 TNDNQTV-TVYVEKASGAPVTWKYVDEBDGNELATPDTLNGKIDAPYOSTAKSISGMWAVKT 704
Qy 328 TFSLDGKDVAISEKIKYQISVNIPLGIADKEG---DANKYKFNLVKXHDALTFDNTVS 384
Db 705 TP-----ANATGVFTDNTQTVY-VYERVDGAPVTWKYVD 738
Qy 365 GEYATALYDQTVIAPEN-----YQTEQA-NGFTVANVPAYIPTL--TPGGLKEV 433
Db 739 G-----DGNELATPDTLNGKIDAPYOSTAKSITGMWTVKTPPANANGVFTDNTQTVTVY 791
Qy 434 YFMHINEKADPTKGFKNBANVDNGHTDQTPPTVEVVTGAKRFIKVGDVTAQALAGAS 493
Db 792 Y-----EKA-----DCAPTVKYVDG-----DGNELAT----- 814
Qy 494 FVVRDQNSDPTANYLKIDETTKAAT-----W--VKTKAATFTTTADGLVDITGLKYST 545
Db 815 -----PDTLNG-KLD-TSYAATAKNLSGMKLTATPANATGVFTTDAQTVTVVYAKQED 865
Qy 546 YLEETVAPDDYVLLTNRIEFVNVNEOSYGTENTVSPKVPKNKHGTLPSGKG----I 601
Db 866 DPKEDKTPPSNTQPDCKTTIKINENKPTSK----PTTIKQTK--LPKTDGDMQESIL 919
Qy 602 VYVLGSGAVLLIAGVYFARR 623
Db 920 FELIGTCFVLL--GIYSVSKK 938
```

Search completed: July 6, 2006, 02:37:23
Job time : 47 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 02:28:16 ; Search time 301 Seconds

(without alignments)
1926.860 Million cell updates/sec

Title: US-10-661-809a-13

Perfect score: 3225

Sequence: 1 MKQLKKWWTSTLLILPL.....GAVLLIAGVFPARRKXENA 627

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3202	99.3	625	2	083618 ENTEROCCU
2	2446	75.8	625	2	03YAD3 ENTEROCCU
3	458	14.2	507	2	048707 LACTOBACILL
4	438.5	13.6	476	2	083619 ENTEROCCU
5	423	13.1	658	2	03Y0Y4 ENTEROCCU
6	395	12.2	473	2	03YAD2 ENTEROCCU
7	355.5	11.0	565	2	03YAC9 ENTEROCCU
8	355	11.0	554	2	03DEE5 STREPTOCOCC
9	355	11.0	554	2	03K250 STREPTOCOCC
10	352	10.9	554	2	03DB92 STREPTOCOCC
11	352	10.9	554	2	08E089 STREPTOCOCC
12	352	10.9	554	2	08E089 STREPTOCOCC
13	345	10.7	693	2	03YAD2 ENTEROCCU
14	337.5	10.5	674	2	03DBD7 STREPTOCOCC
15	334	10.4	502	2	03DBE1 STREPTOCOCC
16	334	10.4	502	2	03K0A5 STREPTOCOCC
17	334	10.4	502	2	08A4A1 STREPTOCOCC
18	333	10.3	705	2	03DV92 STREPTOCOCC
19	333	10.3	705	2	08DYR6 STREPTOCOCC
20	329.5	10.2	674	2	08B4C4 STREPTOCOCC
21	323	10.0	525	2	08B6G6 BIFIDOBACTE
22	296.5	9.2	665	2	0978C2 STREPTOCOCC
23	296	9.2	1530	2	04E006 LISTERIA MO
24	296	9.2	1530	2	08Y479 LISTERIA MO
25	292	9.1	564	2	072XP4 BACILLUS CE
26	292	9.1	564	2	03DMP5 STREPTOCOCC
27	282.5	8.8	554	2	04MW04 BACILLUS CE
28	279.5	8.7	553	2	081D71 BACILLUS CE
29	277	8.6	323	2	03DQW0 STREPTOCOCC
30	276.5	8.6	454	2	02PB82 STREPTOCOCC
31	276.5	8.6	490	2	06NK05 STREPTOCOCC

ALIGNMENTS

RESULT 1	ID	083618 ENTEROCCU	PRELIMINARY	PRT	625 AA.
AC	083618	ENTEROCCU	ENTEROCCU	ENTEROCCU	ENTEROCCU
DT	01-JUN-2003	Integrated into UniProtKB/TrEMBL.			
DT	01-JUN-2003	sequence version 1.			
DT	07-FEB-2006	entry version 12.			
DE	Cell wall surface anchor family protein.				
GN	OrderedLocustName=EF1093; ORFName=EF1093;				
OS	Enterococcus faecalis (Streptococcus faecalis).				
OC	Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.				
OX	NCBI_TaxId=1351;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).				
RC	STRAIN=V583 / ATCC 700802;				
RX	MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;				
RA	Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seebach R.,				
RA	Read T.D., Fouts D.B., Bisen J.A., Gill S.R., Heidelberg J.F.,				
RA	Tetzelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,				
RA	Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,				
RA	Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,				
RA	Khouri H.M., Uetzerback T.R., Radune D., Ketchum K.A., Dougherty B.A.,				
RA	Frazer C.M.,				
RT	"Role of mobile DNA in the evolution of vancomycin-resistant				
RT	Enterococcus faecalis."				
RL	Science 299:2011-2074(2003).				
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms				
CC	Distributed under the Creative Commons Attribution-NonCommercial				
CC	License				
DR	EMBL; AE016830; AM080893.1; -; Genomic DNA.				
DR	TIGR; EF1093;				
DR	BioCyc; EFAB226185:EF1093-MONOMER; -				
DR	GO; GO:0009986; Cell surface; IEA.				
DR	InterPro; IPR008454; Cna B.				
DR	InterPro; IPR001899; Gram_pos_anchor.				
DR	Pfam; PF05738; Cna B; 3.				
DR	Pfam; PF00746; Gram_pos_anchor; 1.				
DR	TIGRfam; TIGR01167; LfXTG_anchor; 1.				
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1.				
KW	Cell wall; Complete proteome; Peptidoglycan-anchor.				
SO	SEQUENCE 625 AA; 67989 MW; EB97450B111F0DE CRC64;				
Query Match	99.3%	Score 3202;	DB 2;	Length 625;	
Best Local Similarity	99.7%	Pred. No. 1.6e-161;			
Matches 625;	Conservative 0;	Mismatches 0;	Indels 2;	Gaps 1;	
QY	1	MKQLKKWWTSTLLILPLFTSVLGTTTFAEENGSAQVYHKKWTDLDPPLIONG	60		
DB	1	MQQLKKWWTSTLLILPLFTSVLGTTTFAEENGSAQVYHKKWTDLDPPLIONG	60		
QY	61	KEMSEPDQYGLADYVFSYVNTNEFYBORAAGASVDAKQAVOSITPEKPYAOGTTDAN	120		
DB	61	KEMSEPDQYGLADYVFSYVNTNEFYBORAAGASVDAKQAVOSITPEKPYAOGTTDAN	120		

121 GNVTVQLPKKNGKDAVYTIKEEPKRGVVAATNMVAPVYEMIKOTDGSYKGTIELAV 180
121 GNVTVQLPKKNGKDAVYTIKEEPKRGVVAATNMVAPVYEMIKOTDGSYKGTIELAV 180
181 VHIYPRNVANDGSL-HKKVGTAEENGNGAEFVSKSGSPQVYKYGVDGLYTWTT 240
181 VHIYPRNVANDGSL-HKKVGTAEENGNGAEFVSKSGSPQVYKYGVDGLYTWTT 238
241 DKEQAKRPTGSKSYEIGENDPTEAENGELTVKMLEVSGSYLLEEVKAPNNMELLENQK 300
241 DKEQAKRPTGSKSYEIGENDPTEAENGELTVKMLEVSGSYLLEEVKAPNNMELLENQK 298
301 TPEFTLEANNQTPVEKTVKNDTSKVDKTPSLDGDVAIGEKIKYOISVNIPLGIADKEGD 360
299 TPEFTLEANNQTPVEKTVKNDTSKVDKTPSLDGDVAIGEKIKYOISVNIPLGIADKEGD 358
361 ANKRYKFNVLVDKHAALTFDNTVSGEYAYALYDGDVIAPENYQVTEQANGFTVAVNPAY 420
359 ANKRYKFNVLVDKHAALTFDNTVSGEYAYALYDGDVIAPENYQVTEQANGFTVAVNPAY 418
421 IPTLPFGTGLKRVYFMHLEKADPTKGFKNENAVNGHDDQTPPTVYVYVGGKRPYKND 480
419 IPTLPFGTGLKRVYFMHLEKADPTKGFKNENAVNGHDDQTPPTVYVYVGGKRPYKND 478
481 GGVATQALAGASPVVRDQNSDPTANYLKIDETTKATVYKTAETFTTTADGLVDITG 540
479 GGVATQALAGASPVVRDQNSDPTANYLKIDETTKATVYKTAETFTTTADGLVDITG 538
541 LKGYTYLLEETVAPDDYVLLTNRIEFVNVESGYTTEMLVSPKVPNKGKGLPSTGKG 600
539 LKGYTYLLEETVAPDDYVLLTNRIEFVNVESGYTTEMLVSPKVPNKGKGLPSTGKG 598
601 IYVYLGAGVLLLIAGVYFARRKENA 627
599 IYVYLGAGVLLLIAGVYFARRKENA 625

RESULT 2
Q3Y3D3_ENTFC PRELIMINARY; PRT; 625 AA.
ID Q3Y3D3_ENTFC PRELIMINARY; PRT; 625 AA.
AC Q3Y3D3;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Surface protein from Gram-positive cocci, anchor region precursor.
GN ORFNames=EfaeBPAF_2569;
OS Enterococcus faecium DO.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=333849;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DO;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RT Hammon N., Israeli S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Enterococcus faecium
DO.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DO;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Lamer F., Lam M.;
RT "Annotation of the draft genome assembly of Enterococcus faecium DO.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DO;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RT Hammon N., Israeli S., Pitluck S., Richardson P.;
RT Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
CC EMBL: AAK0300002; EAN11034.1; -? Genomic_DNA.
DR GO: 0009986; C: cell surface; IEA.
DR GO: 0005618; C: cell wall; IEA.
DR InterPro: IPR008454; Cna_B.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRPFAM: TIGR01167; LpxTG_anchor; 1.
DR PROSITE: PS00847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Signal.
FT SIGNAL 1
FT 33 Potential.
SQ SEQUENCE 625 AA; 68172 MW; D4653B1755F241A2 CRC64;
Query Match 75.8%; Score 2446; DB 2; Length 625;
Best Local Similarity 74.2%; Pred. No. 2,1e-121;
Matches 462; Conservative 64; Mismatches 95; Indels 2; Gaps 1;
1 MKQLKKVYVSTLLIPLFTSVLGTTAPAEENGESQVLVHKKKMTDLPDPLIONSG 60
1 MKGHKKVIMVGLFLLPLPLTNSFGAKKVFABET-AAQVLIHKKKMTDLPDPLIONSG 58
61 KEMSEFDKQGLADYPTGSYVNTNFEYQRAAGASVDAKAQVQSLTRGKPAQGTTPAN 120
59 KEMSEFDYQGLADISFSYVNTQEFYQORDGASVDAKAQVQSLTRGKPAQGTTPAD 118
121 GNVTVQLPKKNGKDAVYTIKEEPKRGVVAATNMVAPVYEMIKOTDGSYKGTIELAV 180
119 GNVTVQLPKKNGKDAVYTIKEEPKRGVVAATNMVAPVYEMIKOTDGSYKGTIELAV 178
181 VHIYPRNVANDGSL-HKKVGTAEENGNGAEFVSKSGSPQVYKYGVDGLYTWTT 240
179 IHLVKNVTVGNDGTLKTKTIGTAENBALNGAEFISKEGTPSVYKIQSVVDGLYTWTT 238
241 DKEQAKRPTGSKSYEIGENDPTEAENGELTVKMLEVSGSYLLEEVKAPNNMELLENQK 300
239 DQTKAKHPTGHSYDIGNDFAEASIEGQLLVNLEVGKINLEEVKAPNNMELLENQK 298
301 TPEFTLEANNQTPVEKTVKNDTSKVDKTPSLDGDVAIGEKIKYOISVNIPLGIADKEGD 360
299 TPEFTLEANNQTPVEKTVKNDTSKVDKTPSLDGDVAIGEKIKYOISVNIPLGIADKEGD 358
361 ANKRYKFNVLVDKHAALTFDNTVSGEYAYALYDGDVIAPENYQVTEQANGFTVAVNPAY 420
359 QNKYTFPLKIDTHDAALTFDNDSSGYAYALYDGNKEIDPVNYSVTEQDGTFSVDPMY 418
421 IPTLPFGTGLKRVYFMHLEKADPTKGFKNENAVNGHDDQTPPTVYVYVGGKRPYKND 480
419 IPTLPFGTGLKRVYFMHLEKADPTKGFKNENAVNGHDDQTPPTVYVYVGGKRPYKND 478
481 GGVATQALAGASPVVRDQNSDPTANYLKIDETTKATVYKTAETFTTTADGLVDITG 540
479 GGVATQALAGASPVVRDQNSDPTANYLKIDETTKATVYKTAETFTTTADGLVDITG 538
541 LKGYTYLLEETVAPDDYVLLTNRIEFVNVESGYTTEMLVSPKVPNKGKGLPSTGKG 600
539 LKGYTYLLEETVAPDDYVLLTNRIEFVNVESGYTTEMLVSPKVPNKGKGLPSTGKG 598
601 IYVYLGAGVLLLIAGVYFARRK 623
599 IYVYLGAGVLLLIAGVYFARRK 621

RESULT 3
Q48707_LACLB PRELIMINARY; PRT; 507 AA.
ID Q48707_LACLB PRELIMINARY; PRT; 507 AA.
AC Q48707;

QY 297 NOTKPTLEANNQTEVKTAKNDTSKVDKTTPLSDGKDAVIGEKIKYOISVNIPLGIAD 356
 DB 188 -----PPELKTIVLDKQGGFQ-----GEMINQTLTQIIPANILG 221
 QY 357 KEGDANKYKFNLDVKHDAALTFDNTSGEYAVALYDGTJAPENYQTEQANGFTYAV 416
 DB 222 -----YQEFRLSDKADTTLTL-----LPESIEVKVAKGVTTT--GYTLTQKHGFTLIDF 268
 QY 417 NPAYIFTLPGGTGLKFVYFMHLNEKADPRTKFNEMN--VDNGHTDDQTPPTVEVTVGGR 475
 DB 269 SIKDQONRA-NQMTWVSQMRLEKTAEPDTAINNEQOLVTDKHT--LTKRATVRRGGS 324
 QY 476 FIKVDGVTATQALAGASFVVDQNSDTANVLIKIDETTKAATWTKAEATFTTTADGL 535
 DB 325 FVKVDSE-NAKITLBEAVFIVKNGQGEYIN-----ETANGYMQKEKALAKKFTSNQGE 378
 QY 536 VDTGLKGTYYLLEETVADPDYVLLTNRIEFVNEQSYGTENLVSEKVPN--KHKGT 592
 DB 379 FSVKGLKQGYPLEEISAPKGYILNQTEIPFTVKGKSYATNGQRTAPFLHVNKKYKESGF 438
 QY 593 LPSTG-GKGIYVYLGSAVLLLIAGV-----YFARRKE 625
 DB 439 LPKTBERSITWLTAG---LIIIGVAVIMLFYQKQKGE 474

RESULT 5

Q3Y0Y4_ENTFC

Q3Y0Y4_ENTFC PRELIMINARY; PRT; 658 AA.

AC Q3Y0Y4; 11-OCT-2005, integrated into UniProtKB/TrEMBL.

DT 11-OCT-2005, sequence version 1.

DT 07-FEB-2006, entry version 3.

DE Surface protein from Gram-positive cocci, anchor region precursor.

GN ORFNames=BfaeRART_1904;

OS Enterococcus faecium DO.

OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.

OX NCBI_Taxid=333849;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=DO;

RG US DOE Joint Genome Institute (JGI-PGF); Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,

RA Hammon N., Israeli S., Pritchuck S., Richardson P.;

RT "Sequencing of the draft genome and assembly of Enterococcus faecium

DO."; Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=DO;

RG US DOE Joint Genome Institute (JGI-ORNL);

RA Larimer F., Land M.;

RT "Annotation of the draft genome assembly of Enterococcus faecium DO."; Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=DO;

RG US DOE Joint Genome Institute (JGI-PGF);

RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,

RA Hammon N., Israeli S., Pritchuck S., Richardson P.;

RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms

CC Distributed under the Creative Commons Attribution-NonCommercial license

CC EMBL; AAK0300014; EMBL0156.1; ? Genomic DNA.

DR GO; GO:0009986; C:cell surface; IBA.

DR InterPro; IPR008454; Cna_B.

DR InterPro; IPR001899; Gram_pos_anchor.

DR Pfam; PF05738; Cna_B_2.

DR TIGRfams; TIGR01167; LPXTG_anchor; 1.

KW Signal. 1 32 Potential.
 FT SIGNAL. 658 AA; 71457 MW; F540948280207BA CRC64;
 SQ SEQUENCE

Query Match 13.1%; Score 423; DB 2; Length 658;
 Best Local Similarity 25.7%; Pred. No. 4,7e-14;

Matches 188; Conservative 101; Mismatches 254; Indels 188; Gaps 38;

QY 2 KQKKVWTVSTLLILFLPSVLTGTTAPAE-----NGESAQVVIHKKMTDLPDP- 55
 DB 5 RYKKAAGLIGMLIGLIACPGQY---RAYADVTQKTPPEKVNITVH-KLMTYDGTQJLN 59
 QY 56 -----IQNSGKMSF-----DKYQ--GLADYFYSYNTNEY---EQRAGASVDAKQ 101
 DB 60 VDIDIKNDGTHDDYPPGVTKYNNKADYGVDEFITGNITDQVLPREDSDLTNAKDELYK 119
 QY 102 AVQSILPGKPVAQGT-----DANGVT-VQLPKQNGKDAVTTIK-EPEGVVA-- 150
 DB 120 DVEDGSGNSEYKNAKNTKTSVADNGEITFPADQPAVYNSKGNVYVESKSAAGLVQK 179
 QY 151 ATNMYVAPPVYEMIKQDGSYKYEELAVVHIYKRVVANDGSLHKV--GTAENE- 206
 DB 180 AKPMVVIAPM-----TNT--SGFLKDHLHYKNTVSK-LSFELTKFGDDGTAQSKQ 229
 QY 207 -GLNGAEFVTSKESGSPGVKXYIQGVKQGLYTTTDEKQAKRFTYKSYEIGENDTEAE 265
 DB 230 TPLKGAKEFLYKGE-----PGKGTLDG-----LVS 255
 QY 266 NOTGELTVKNTLEFVGYIIIEYKAPNNAELIENQTTPT-----IEANQTPVEK 315
 DB 256 DDQGLTTLTDLTLGKYFEVVP---SEVVGSDPEFPADQYLLGADARNDAHNLTEI 311
 QY 316 TVKNDTSK-----VDKTPSLDGKD--VAIGEKIKYOISVNIPLGIADKEGDA 361
 DB 312 TNDGVSLSLKASVYNYKAPVLDKTYTNGTQGHSGQIGDAVAVYQSTHIFPDIA---GGA 368
 QY 362 NKCYKFNLDVKHDAALTFDNTSGEYAVALYD-GDT-----VI 398
 DB 369 D-----GITVNGVSETPSPVSFPMGDTAQLGSLYAKANIKVTNKGSGV 415
 QY 399 APEN--YQTEQANGFTV--AVNPAYIPTLP--GGLTKFVYFMHLNEKADPTGPFQNE 451
 DB 416 LKENIDYKIQNSSENFVIDFIVNNGQVSDPYASLHGDLQMTNMYVNDVSAVANPLTNS 475
 QY 452 AN-VDNGHTDDO---TPPTVEVTVGKRFIVDGDVAT-----QALGASFVVDQNSD 502
 DB 476 VDFVTNNPFPNGEHEHEKTKADVVTYGAFLKVDSGLFGTGIKATPLGSAEPAK--NAE 533
 QY 503 TANYLKIDET---TKAATWTKAEATFTTTADGLVDITGLKGYTLYLEETVADPDYV 558
 DB 534 GKYYGLVDTDKDGKCAVWVDVANAAIILKSDKEGHFEITGLTEGYSLEETKAPENYQ 593
 QY 559 LITNRIEFVNAEQSYGTENLVSPKVPYKHKGTLPSTGKGIIYY-----LQSGAVL 611
 DB 594 KLTKESIFKVDKDY-KEENRIT---IKNNQASVPMTGSGNGFQTYVLISCLLIDAGA-- 647
 QY 612 LLIAGYFARR 622
 DB 648 -LSAVVYFKKK 657

RESULT 6

Q3Y3D2_ENTFC

Q3Y3D2_ENTFC PRELIMINARY; PRT; 473 AA.

AC Q3Y3D2; 11-OCT-2005, integrated into UniProtKB/TrEMBL.

DT 11-OCT-2005, sequence version 1.

DT 07-FEB-2006, entry version 3.

DE Surface protein from Gram-positive cocci, anchor region; Cna B-type

GN ORFNames=BfaeRART_2570;

OS Enterococcus faecium DO.

OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.

Best Local Similarity 22.7%; Pred. No. 1.5e-10;
Matches 159; Conservative 98; Mismatches 221; Indels 223; Gaps 30;

QY 5 KKWVTVSTLLILPLFVSLGTTAFPAENGESAO-----LVHKKMTDLPDLION 58
DB 3 KKMIVTGSVLLILVAGIGIFGNGKLVKAE--EVAQKPEEVITLHKKGFSSVPEER -PN 59
QY 59 SKKEMSEPKYOGGLDVPFSIVNTNEFE-----QRAAGASVDAKQAVOS----- 105
DB 60 SLVSTDBGE-ENIGEVDPDLPVTEVYDILRDNPLTEREDGLASALIEWIQKRTTE 118
QY 106 -----LTPGKPVAGTTDANGV---TVQLPKK-QNGKDAVYTIKE--EPKEGVAAAT 152
DB 119 SWFLKRLIL---SIDKQTNEAGAVFSIVQYTEAPSRDKVYLELFETYSAPHSRLAS 175
QY 153 NMVAVAFPVYEMIKQIDGSY-----KYTEELAVHHIYKPVNVANDGSLHKVKGTAEN 205
DB 176 PAVVMMVPM-MEDMVDGVWDGSTWKDVTNTD---VHLVPKN----- 212
QY 206 EGLNGAEFVSKSESGPGTVKVIYIGVKGGLYMTTDDKEQAKRFTGKSYEIGENDPTEAE 265
DB 213 -----EIRBAD----- 218
QY 266 NGTGLVKNLEVGSIYIEEVKAPNNABLIENQTKPTTEANNQTPVEKTVKNDTSKYD 325
DB 219 -----KQNVBESDLRQVTIIN-----EAGEQETISYI----- 246
QY 326 KTTPLSDGDVAIGEKIKQIISVNIPLGIADKEGDANKVK-PNLVDKDAALTPDNVTS 384
DB 247 -----DLERKTAASYITAPIPYFIDVSLVENSASAVIKNYKITDPTVGLTY----- 292
QY 385 GEYAAVALYD-----GDTVIAP-ENVQYTEOANGFTVAV-----NPAVYIPLT-----TP 426
DB 293 -----YDQELVRAGETIITLKGDYIYEVVNSNGVVTIILTBENGVAIVDLGLADAR 345
QY 427 GGTLEFVYFMHLEKADPTKGFKNENAVDNGHTD-----DQTPPEVVTGSKRFLKVDG 481
DB 346 GGDLTITVYLKSTLEADDFHNNFVAVIEIGRNDDEYEGVEPPEKVTITGGRKFEKY-- 403
QY 482 DVTATQALAGASPVVARDNSDTANY-----LKIDETTKAATVTKAETTTTAA 532
DB 404 DASSSELMLDARFEL--WNEDRSEVAIFYKGSPLAVYESGADRLEWATSGATEFVADG 461
QY 533 DELVNTITGKYGTYYLEETVAPDDYVLLTNR---IFVNVNEOSYGTTEMLV-----SPK 584
DB 462 NQYFEVQGLDYGTYQMKETMAPEGVLPFGAALFTFIISYGSYNELIOIVGVENGPFR 521
QY 565 VNNKKGTLPTSGKGIYVYLGSAVLLLIAGVYFARRKE 625
DB 522 VNNMKRGLSLPATGNGGLAFLILIGISLMI--GAYSWYRKS 560

RESULT 8
Q3DEF5_STRAG PRELIMINARY; PRT; 554 AA.
AC Q3DEF5;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Cell wall surface anchor family protein.
GN ORFNames=SAM_0661;
OS Streptococcus agalactiae CUB11.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
NCBI_TaxID=342617;
RN NCB1
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CUB11;
RX PubMed=16172379; DOI=10.1073/pnas.0506758102;
RA Terehan H., Masigiani V., Cielewicz M.J., Donati C., Medini D.,
Ward N.L., Anguioi S.V., Crabtree J., Jones A.L., Durkin A.S.,
DeBoy R.T., Davidson T.M., Mora M., Scaletti M., Margalit Y. Ros I.,
Peterson J.D., Hauser C.R., Sundaram J.P., Nelson W.C., Madupu R.,

RA Brinkac L.M., Dodson R.J., Rosevitz M.J., Sullivan S.A.,
RA Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N.,
RA Khouri H., Radune D., Dimitrov G., Mackins K., O'Connor K.J.,
RA Smith S., Utterback T.R., White O., Rubens C.B., Grandi G.,
RA Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R.,
RA Frazer C.M.;
RT "Genome analysis of multiple pathogenic isolates of Streptococcus
RT agalactiae: implications for the microbial 'pan-genome'." ;
RL Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL; AAU00100012; EAO73672.1; - Genomic DNA.
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0009986; C:cell surface; IEA.
SQ SEQUENCE 554 AA; 60150 MW; 64498CB80A075ED1 CRC64;

Query Match 11.0%; Score 355; DB 2; Length 554;
Best Local Similarity 27.3%; Pred. No. 1.5e-10;
Matches 164; Conservative 67; Mismatches 225; Indels 144; Gaps 29;

QY 125 VOLPKKQKDAVYTI-----KEPKEGVAAATNMVAVPYVE--MIKQDGSYKYGTBEL 178
DB 1 MKLSKLLFSAVALITVAGSTVEPAQFATGSIYRAAEVGERAKTTVNIYKLQADS- 59
QY 179 AVVHHYKPVNVANDGSLHKVKKVGTAEENGAEFVSKSESGPGTVKVIYIGVKGGLYTW 238
DB 60 -----YKSEIISNGGI-----ENKQGE--VISNYAKLGDVNGVGLQGVQPKRYKV 101
QY 229 TTD--KEQAKRIT-----GKSYEIGENDFTEANGTGLVKNLEVGSI---LE 284
DB 102 KTDISVDELKLTVEADAKVGTILBEGVS-LPQKTAQG-LVVDALDSKNVAYLYVE 159
QY 285 EVK-APNN-----AEL-LENQTKPTTEANNQTPVEKTVKNDTSKYDKTTPSLD 332
DB 160 DLKNSPNTKAYAVPVLELPLVANSSTGFLSEIN---IYKVVVTDEPKTDKQVKGL- 215
QY 333 GKQVA---IGEKIKQIISVNIPLGIADKEGDANKVYKRNLDVKHDAALTFDNV----- 382
DB 216 GQDDAGYITGEFFKFKFLKSTIPANLGD-----TEKEEITDKFADGLTYKSVGKIKGS 268
QY 383 -TSGEYAAVALYDGVV-----IAPENY-QVTEQANGFTVAVNPAVYIPLT---PG 427
DB 269 KTLNDEHTTIDEPVNDONTLKITFKPEKFEIALLKMTLVKQDALDQATNTDQA 328
QY 428 GTLKEFVYFMHLEKADPTKGFKNENAVDNGHTDQ-----PPEVVTGSKRFLKV 479
DB 329 AFLBIPVASTINEKAVLGAIENTPELOYDHTPPKADNPKEPNPRKEVHTGGRFVK 388
QY 480 DGDVTAIQLAGASPVVARDNSDTANYLKIDETTKAATVTKAATV----- 527
DB 389 DS--TETDTLGAER-----DLASDQ--AVKWTALIKANTKNYIAGEAVTG 434
QY 528 -----FTTTADGLVDITGLKYG-----TYYLEETVAPDDYVLLTNRIEFVNEOSY 573
DB 435 QPIKLSHTDGTFFELKGLAYVANDABGATVYLLKETKAPBGVYIIPKKEIFYSQSY 494
QY 574 GT--TENVV-----SPKVPNKGKGLTPSTGKGIYVYLGSAVLLLIAGVYFARRKEN 626
DB 495 NTKPFDITVDSADAPPTIKNNKRSIPNTGIGTAIFVALGAAVMAFVKGKMRRTKDN 554

RESULT 9
Q3K250_STRAI PRELIMINARY; PRT; 554 AA.
AC Q3K250;
DT 08-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 08-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Cell wall surface anchor family protein.

GN OrderedLocusNames=SAK_0776; ORFNames=SAK_0776;
 OS Streptococcus agalactiae serotype Ia.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=355315;
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX STRAIN=909 / ATCC 27591 / Serotype Ia;
 PubMed=16172379; DOI=10.1073/pnas.0506758102;
 RA Tettein H., Masignani V., Cieslewicz M.J., Donati C., Medini D.,
 Ward N.L., Anguoli S.V., Crabtree J., Jones A.L., Durkin A.S.,
 DeBoy R.T., Davidson T.M., Mora M., Scarselli M., Margarit Y Ros I.,
 Peterson J.D., Hauser C.R., Sundaram J.P., Nelson W.C., Madupu R.,
 Brinkac L.M., Dodson R.J., Rosovitz M.J., Sullivan S.A., Zhou L., Zafar N.,
 Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N.,
 Khoult H., Radune D., Dimitrov G., Watkins K., O'Connor K.J.,
 Smith S., Uetebach T.R., White O., Rubens C.B., Grandi G.,
 Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R.,
 Fraser C.M.;
 RA "Genome analysis of multiple pathogenic isolates of Streptococcus
 agalactiae: implications for the microbial 'pan-genome'.";
 RL Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955(2005).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL: CP000114; AAA44713.1; -; Genomic_DNA.
 DR TIGR: SAK_0776; -;
 DR GO: GO:0009986; C:cell surface; IEA.
 DR InterPro: IPR008454; Cna_B.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR Pfam: PF00746; Gram_pos_anchor; 1.
 DR TIGRfam: TIGR01167; LpxTG_anchor; 1.
 KW Complete proteome.
 SQ SEQUENCE 554 AA; 60150 MW; 64498CB80A075BD1 CRC64;

Query Match 11.0%; Score 355; DB 2; Length 554;
 Best Local Similarity 27.3%; Pred. No. 1.5e-10;
 Matches 164; Conservative 67; Mismatches 225; Indels 144; Gaps 29;

125 VOLPKKQNGDAVYTI-----KEEPKEGVAAATMVAAPVYE--MIKOTDGSYKYTEEL 178
 1 MKLSKKLFSANLTIIVAGSTVEPVAQPATGMSIVAAASQSRPAKTTVNIYKQADS- 59
 179 AVVHIYPRKVVANDSLHVKVGTAEENGANGAEFVYSSESGPGTVKXIQQVKDGLYTM 238
 60 -----YKSEITSNGGI-----ENKDG--VISNYAKLGDVVKGLQGVQFRKRYK 101

239 TTD--KEQAKRPT-----GKSYEIGENDPTEANGTGLTVKNLEVGSI-----LE 284
 102 KTDISVDELKGLTVEADAKVGTILEEVS--LPQRTNQG--LVVADLSKSNVRYLYVE 159
 285 EVK-APNN-----AEL--IENQTKPTTEANNQTPVEKTVKNDTSKYDKTPSID 332
 160 DLKNSNSNTTKAYAVPVELEPVANSTGGLSEIN---IYKNNVTDEPKDKVKKL- 215
 333 GKDVA---IGEKIKQIYSNIPGLIADKEGDKANKYKFNLVKHAALTTFDNY----- 382
 216 GODDAGYTTIGBEFKFPLKSTIPANLGD-----YEKFEITDKFADGLTYKSVGKIKIGS 268
 383 -TSGEVAVLALYDGT-----IAPENY-QVTEQANGFTVAVPAIYPLT-----PG 427
 269 KTLNDEHTTIDPEYDNONTKITTKPEKFEKIALKMTLVKQODADKATANTDDA 328
 428 GTLKFPYFPAHLEKADPTGKFKNEAVNDGHTDQ-----PTVEVVTGSKKFIKY 479
 329 APLEIVASTINKEKAVLGKAIENTFELQYDHTPDKADNPESNPPEKPEVHGGKFPVK 388
 480 DGDVTAFTQALAGASFVVRDONSPTANYLKIDETTKAATVVKAKAFT----- 527
 389 DS--TETQTLGGAEF-----DLASDGT--AVKWTDALIKANTKNVYIAGEAVYTG 434

QY 528 -----FTTADGLVDITGLKYG-----TYYEEFVAPDDYVLLTNRIEFVANEQSY 573
 DB 435 QPKLKSHTDGFREIFGLAYVADANAEATVAYTKAKETAPBEYVLPDKEIETFGQTSY 494
 QY 574 GT--TENLY-----SEKVPNPKKGTLPSTGKGIYVYLGSGVLLIAGVFAARRKEN 626
 DB 495 NTFEPTDITVDSADATPDITKNNRPSIPNTGGIGTALFVALGAANVAFVAKGMRRTKDN 554

RESULT 10
 Q3DB92 STRAG PRELIMINARY; PRT; 554 AA.
 ID Q3DB92;
 AC Q3DB92;
 DT 22-NOV-2005. Integrated into UniProtKB/TrEMBL.
 DT 22-NOV-2005. sequence version 1.
 DT 07-FEB-2006. entry version 4.
 DE Cell wall surface anchor family protein.
 GN ORFNames=SAN_0698;
 OS Streptococcus agalactiae COH1.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=342616;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=COH1;
 RX PubMed=16172379; DOI=10.1073/pnas.0506758102;
 RA Tettein H., Masignani V., Cieslewicz M.J., Donati C., Medini D.,
 Ward N.L., Anguoli S.V., Crabtree J., Jones A.L., Durkin A.S.,
 DeBoy R.T., Davidson T.M., Mora M., Scarselli M., Margarit Y Ros I.,
 Peterson J.D., Hauser C.R., Sundaram J.P., Nelson W.C., Madupu R.,
 Brinkac L.M., Dodson R.J., Rosovitz M.J., Sullivan S.A., Zhou L., Zafar N.,
 Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N.,
 Khoult H., Radune D., Dimitrov G., Watkins K., O'Connor K.J.,
 Smith S., Uetebach T.R., White O., Rubens C.B., Grandi G.,
 Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R.,
 Fraser C.M.;
 RA "Genome analysis of multiple pathogenic isolates of Streptococcus
 agalactiae: implications for the microbial 'pan-genome'.";
 RL Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955(2005).
 CC -!- CAUTION: The sequence shown here is derived from an
 preliminary data.
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL: AAJ0100002; BA076719.1; -; Genomic_DNA.
 DR GO: GO:0009986; C:cell surface; IEA.
 SQ SEQUENCE 554 AA; 60168 MW; 64498D4B49729AD1 CRC64;

Query Match 10.9%; Score 352; DB 2; Length 554;
 Best Local Similarity 27.2%; Pred. No. 2.2e-10;
 Matches 163; Conservative 68; Mismatches 225; Indels 144; Gaps 29;

125 VOLPKKQNGDAVYTI-----KEEPKEGVAAATMVAAPVYE--MIKOTDGSYKYTEEL 178
 1 MKLSKKLFSANLTIIVAGSTVEPVAQPATGMSIVAAASQSRPAKTTVNIYKQADS- 59
 179 AVVHIYPRKVVANDSLHVKVGTAEENGANGAEFVYSSESGPGTVKXIQQVKDGLYTM 238
 60 -----YKSEITSNGGI-----ENKDG--VISNYAKLGDVVKGLQGVQFRKRYK 101

239 TTD--KEQAKRPT-----GKSYEIGENDPTEANGTGLTVKNLEVGSI-----LE 284
 102 KTDISVDELKGLTVEADAKVGTILEEVS--LPQRTNQG--LVVADLSKSNVRYLYVE 159
 285 EVK-APNN-----AEL--IENQTKPTTEANNQTPVEKTVKNDTSKYDKTPSID 332
 160 DLKNSNSNTTKAYAVPVELEPVANSTGGLSEIN---IYKNNVTDEPKDKVKKL- 215
 333 GKDVA---IGEKIKQIYSNIPGLIADKEGDKANKYKFNLVKHAALTTFDNY----- 382
 216 GODDAGYTTIGBEFKFPLKSTIPANLGD-----YEKFEITDKFADGLTYKSVGKIKIGS 268

Db 650 ORENK-KVTIPOTGIGITLFTIIGLSIMLGAVIMKRROSEA 693

RESULT 14

Q3DERS STRAG PRELIMINARY; PRT; 674 AA.

AC Q3DERS;

DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.

DT 22-NOV-2005, sequence version 1.

DT 07-FEB-2006, entry version 5.

DE Cell wall surface anchor family protein.

ORFNames=SAM_1372;

OS Streptococcus agalactiae CUB111.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

NCBI_TaxID=342617;

OX NCBI_TaxID=342617;

RN NUCLEOTIDE SEQUENCE.

RP STRAIN=CUB111;

RC PubMed=16172379; DOI=10.1073/pnas.0506758102;

RA Tettelin H., Maignani V., Cielesiewicz M.J., Donati C., Medini D., Ward N.L., Angiuoli S.V., Crabtree J., Jones A.L., Durkin A.S., DeBoy R.T., Davidson T.M., Mora M., Scarselli M., Margarit y Ros I., Peterson J.D., Hauser C.R., Sundaram J.P., Nelson W.C., Madupu R., Brinkac L.M., Dodson R.J., Rosovitz M.J., Sullivan S.A., Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N., Khouri H., Radune D., Dimitrov G., Watkins K., O'Connor K.J., Smith S., Uterback T.R., White O., Rubens C.E., Grandi G., Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R., Frazer C.M.;

RT "Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial 'pan-genome'.";

RL Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955(2005).

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NonDerivs License

CC -----

DR EMBL, AAJ001000018; EAO73402.1; -, Genomic_DNA.

DR GO, GO:0009986; C:cell surface; IEA.

DR GO, GO:0005618; C:cell wall; IEA.

DR KW Cell wall; Peptidoglycan-anchor.

DR SQ SEQUENCE 674 AA; 73217 MW; 5463FC096BCF0022 CRC64;

Query Match 10.5%; Score 337.5; DB 2; Length 674;

Best Local Similarity 25.1%; Pred. No. 1.7e-09;

Matches 192; Conservative 78; Mismatches 267; Indels 227; Gaps 39;

QY 1 MKQLKQWVTVSTLLILPLFTSVIGTTTAPAEENGESQVLVHKKKMTDLDPDLION-- 58

DB 1 MKKIKKLTMTSTLIL--TSLPSVAPAFDD-ATTVTVLHAKVM---PQAAFDNNT 53

QY 59 -----SGKEMSEFDKYQGLADVT-----FSIYNVT-NEFEYGRAGASVDA-- 98

DB 54 EGTGKNDSDYVQKQINDLKSYFGSTDAKEIKGAFVFNERTGKPTEN--GKEVDLIE 111

QY 99 AKQAVQSLTPGKPAVAGTTDANGNV--TYQLPKQNGKDAVTTIIE--PKEGVAAATN 153

DB 112 AKDA-----EGGAVLSGLTKDNGFVFNIAKL-----KGIVQVELKEKSNYDNNGSLIDS 162

QY 154 MVVAPFVYEMIKOTGSGYKGTBEELAVVHIYPRNVAN--DGLSLVKKVGTAEENGANG 210

DB 163 KAVPVKITPLVNNQGVVK-----DAHIYKNTETKQVDKNRPADKLDITDNRKDG 215

QY 211 AEFVSKSESGPQVYKIQGVK-----DGLVTWTTDKQAKRFTGSKSYEL- 256

DB 216 ---VVSATVGD--KKEVIYIGTKLKGSDYKLVMTDSMTKGLTFNNNVNVTLDGEDFVL 270

QY 257 -----GENDPTEAENGTELVY-----KNEVGSYIIEVKAPRNALIE-NQTKIPF 303

DB 271 NYKLVTDGQFRLANATGLAAVAAAADKDVIEIKITYSATVNGSTTVEIPETNDVKLDY 330

QY 304 TIEANNOT-----PVEKTVKNDTSKYVDKTTPLSDGKDAVIGEKIKYQIVSNVPIGIADKEG 359

DB 331 ---GNPTEBSEFPQGTIPANOEIKVI-----KDMAVDGIIT----- 363

QY 360 DANKYVK--FNLVDKDAALTPDNVTSGEYAYALYDGDVTAJAPEN---YQVTEQANGFTV 414

DB 364 DANVAVKAIPTLQEKQTDG--TWNVVASHATKPSRFEHTFTGLDNAKTVRVERVSGYT- 421

QY 415 AVNPAYIPLTPPGTLKTVYFMHLEKADPTKGFQNEANV--DNCHTDQGT---PPTVEV 469

DB 422 ---PEYV-----SFGNGVVTIKNNKNSDPTINSEPKV 453

QY 470 VTGKRFPFKVDGVTATQALAGSFVRDQNSDTN-----YLKIDETT 513

DB 454 VTYGKFPVKT--NQANTERLAGATTLVKKEGKYLARKGAATPAEKAATKAKLADLEV 511

QY 514 KA-----AT-----WVTKA--EATPTTTAD 533

DB 512 KAVNDLTKEQEGQGTALATVDQKQAYNDAPVKANYSEWVADKKADNVKLIISNAG 571

QY 534 GLVDTGKLYGTYVLEETVAPDDYLLNTRIEFPVNBGSY--GTENLV-----SPE 583

DB 572 GQFELTGLDKGTYGLEEQAPAGVATLSGDVFEVTATSYSKGATTDIAIDKGSYKDAQ 631

QY 584 KVPNGHKTLPSGTGKGIYVYVLSGAVLLLAGVYFARRRENA 627

DB 632 QVQNK-KVTIPOTGIGITLFTIIGLSIMLGAVIMKRROSEA 674

RESULT 15

Q3DRT1 STRAG PRELIMINARY; PRT; 502 AA.

AC Q3DRT1;

DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.

DT 22-NOV-2005, sequence version 1.

DT 07-FEB-2006, entry version 4.

DE Cell wall surface anchor family protein, putative.

GN ORFNames=SAN_1518;

OS Streptococcus agalactiae COH1.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

NCBI_TaxID=342616;

RN NCBI_TaxID=342616;

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=COH1;

RC PubMed=16172379; DOI=10.1073/pnas.0506758102;

RA Tettelin H., Maignani V., Cielesiewicz M.J., Donati C., Medini D., Ward N.L., Angiuoli S.V., Crabtree J., Jones A.L., Durkin A.S., DeBoy R.T., Davidson T.M., Mora M., Scarselli M., Margarit y Ros I., Peterson J.D., Hauser C.R., Sundaram J.P., Nelson W.C., Madupu R., Brinkac L.M., Dodson R.J., Rosovitz M.J., Sullivan S.A., Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N., Khouri H., Radune D., Dimitrov G., Watkins K., O'Connor K.J., Smith S., Uterback T.R., White O., Rubens C.E., Grandi G., Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R., Frazer C.M.;

RT "Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial 'pan-genome'.";

RL Proc. Natl. Acad. Sci. U.S.A. 102:13950-13955(2005).

CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NonDerivs License

CC -----

DR EMBL, AAJ001000022; EAO75904.1; -, Genomic_DNA.

DR GO, GO:0009986; C:cell surface; IEA.

DR SQ SEQUENCE 502 AA; 53450 MW; E3B7F5603F6D6192 CRC64;

Query Match 10.4%; Score 334; DB 2; Length 502;

Best Local Similarity 26.3%; Pred. No. 1.7e-09;

Matches 150; Conservative 79; Mismatches 223; Indels 118; Gaps 26;

```
QY 100 KQAVQSL-----TFGRVAGQTTANGNVTVQLPKQNGKDAVYTIKEEPEKV 148
Db 3 KKNIQSLVLVASLAFGNAVSPVPIAFPA--ETGITVQDQKATYKAYRVFPAEIDNAN 60
QY 149 VAATNNVVAFPVYEMIKQTDGSYKGTBELAVVHVPKNVANDSLHVKKVGTAE-NEG 207
Db 61 VSDSNQDGA--SLYIPQGEAEYKASTDPSLSP-----TTTNGGRYVTKQDTASANE- 112
QY 208 LNGAEFVVISKESSPGTVKXIQGVKQGLYTTWTDKQAKRFITGKSYEIGENDPTEAEN- 266
Db 113 -----IATWA-----KSIANTTPV--STYTESND 136
QY 267 GTGELTVKNLEVGSYLLBEVKAPNNABLIENQTKPTPTIANNQTPVEKTVKNDTSKVDK 326
Db 137 GTEVINVS--QYGYVVS--STVNGAVIMVTSVTP-----NATIEK--NTDATWGDG 184
QY 327 TFPSLDKQVAIGEKIKYQISVNIPLGIADKEGDANKYKFNLVDKHDALEFDNYTSGE 386
Db 185 GCKTVQKTYSVGDYVKYITTYKNAVNYHCTE---KVYQYIKDTMPSASVVD-LNEGS- 239
QY 387 YAVVALYDSDTVI-----APENYQVTEQANGFTVAVNPAYIPTLTPOG----- 428
Db 240 YEVITIDGSGNITTLTGSEKATGKYNLENNNFITTI--PWAATNTPGTONGANDD 297
QY 429 -----TLKFVYFPHLNEKADPTKG---FKYEAENVDNQHTDDQTPPYEVYTGKRFI 477
Db 298 FFYKGINITITVYTVGLKSGAKPGSADLPENTNIATINPNTSNDPQOKTVARDQIYIK 357
QY 478 KYDGDVYATQALAGASFFVVRDQNSDTANYLKIDETAKATWVKYKAATFTTTADQLVD 537
Db 358 KIDGSTKA--SLOGAIFVLKNA---TGQFLNPDNTNN-VEW-GTEANATEYTTGADGIIT 410
QY 538 ITGLKXGYIYLBETVAPDDYVLLTNRIBFVNVQSYGTT--ENLVSPKVPNKHGKT-LP 594
Db 411 ITGLKEGTYLVVEKKAPLGYNLLDNSQKVLIGDGAIDTTSNDLVLVNPTEVN-NKGTELP 469
QY 595 STGGKGIYVYVLSGANLULLIAGVYFARRRK 624
Db 470 STGGIGTTFYIIGALVLVIGAGIVLVARRR 499
```

Search completed: July 6, 2006, 02:36:34
Job time : 306 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 02:48:51 ; Search time 166 seconds

(without alignments)
1561.462 Million cell updates/sec

Title: US-10-661-809a-13

Perfect score: 3225
Sequence: 1 MKQLKKWYTVSTLLILPL.....GAVLLINGVPRARRKENA 627

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:*\n2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:*\n3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:*\n4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*\n5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*\n6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3225	100.0	627	3	US-09-071-035-218 Sequence 218, App
2	3225	100.0	627	4	US-10-206-576-218 Sequence 218, App
3	3225	100.0	627	4	US-10-661-809-13 Sequence 13, Appl
4	3225	100.0	627	5	US-10-912-362-218 Sequence 218, App
5	2892	88.7	560	3	US-09-071-035-220 Sequence 220, App
6	2892	88.7	560	4	US-10-206-576-220 Sequence 220, App
7	2892	88.7	560	5	US-10-912-362-220 Sequence 220, App
8	2446	75.8	625	4	US-10-661-809-19 Sequence 19, Appl
9	438.5	13.6	476	4	US-10-661-809-11 Sequence 11, Appl
10	395	12.2	473	4	US-10-661-809-17 Sequence 17, Appl
11	347	10.8	430	3	US-09-071-035-274 Sequence 274, App
12	347	10.8	430	4	US-10-206-576-274 Sequence 274, App
13	347	10.8	430	5	US-10-912-362-274 Sequence 274, App
14	334	10.4	502	4	US-10-333-002-16 Sequence 16, Appl
15	311	9.6	384	3	US-09-071-035-276 Sequence 276, App
16	311	9.6	384	4	US-10-206-576-276 Sequence 276, App
17	311	9.6	384	5	US-10-912-362-276 Sequence 276, App
18	296.5	9.2	665	3	US-09-769-787-127 Sequence 127, App
19	296.5	9.2	665	5	US-10-472-928-778 Sequence 778, App
20	296.5	9.2	665	5	US-10-873-528-127 Sequence 127, App
21	296	9.2	1530	6	US-11-045-004-34 Sequence 34, Appl
22	267	8.3	1612	6	US-11-045-004-38 Sequence 38, Appl
23	266	8.2	1788	4	US-10-282-122A-46664 Sequence 46664, A
24	245.5	7.6	724	4	US-10-282-122A-45795 Sequence 45795, A
25	242.5	7.5	793	4	US-10-282-122A-60689 Sequence 23, Appl
26	242.5	7.5	793	6	US-11-045-004-23 Sequence 23, Appl
27	238.5	7.4	245	4	US-10-333-002-2 Sequence 2, Appl

28	232	7.2	2032	3	US-09-071-035-458 Sequence 458, App
29	232	7.2	2032	3	US-09-071-035-462 Sequence 462, App
30	232	7.2	2032	3	US-09-071-035-466 Sequence 466, App
31	232	7.2	2032	4	US-10-206-576-458 Sequence 458, App
32	232	7.2	2032	4	US-10-206-576-462 Sequence 462, App
33	232	7.2	2032	4	US-10-206-576-466 Sequence 466, App
34	232	7.2	2032	5	US-10-912-362-458 Sequence 458, App
35	232	7.2	2032	5	US-10-912-362-462 Sequence 462, App
36	232	7.2	2032	5	US-10-912-362-466 Sequence 466, App
37	226	7.0	489	4	US-10-282-122A-57763 Sequence 57763, A
38	217.5	6.7	621	3	US-09-071-035-468 Sequence 468, App
39	217.5	6.7	621	4	US-10-206-576-468 Sequence 468, App
40	217.5	6.7	621	5	US-10-912-362-468 Sequence 468, App
41	206	6.4	1103	4	US-10-661-809-9 Sequence 9, Appl
42	202.5	6.3	402	5	US-10-472-928-780 Sequence 780, Appl
43	193.5	6.0	940	6	US-11-045-004-39 Sequence 39, Appl
44	189	5.9	688	3	US-09-071-035-464 Sequence 464, App
45	189	5.9	688	4	US-10-206-576-464 Sequence 464, App

ALIGNMENTS

RESULT 1
US-09-071-035-218
Sequence 218, Application US/09071035
Publication No. US20020045737A1
GENERAL INFORMATION:
APPLICANT: G11 H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
FAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 218:
SEQUENCE CHARACTERISTICS:
LENGTH: 627 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-218
Query Match
Best Local Similarity 100.0%, Pred. No. 9.6e-188;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKQLKKWYTVSTLLILPLFTSVIGTTTAAENGESAAQVIHKKKXTDLPDPLIONSG 60
DB 1 MKQLKKWYTVSTLLILPLFTSVIGTTTAAENGESAAQVIHKKKXTDLPDPLIONSG 60

QY 61 KEMSEFDKXQGLADYTFSIYVNTNEFEQRAAGASVDAKQAVQSLTPGKPVAGTTDAN 120
DB 61 KEMSEFDKXQGLADYTFSIYVNTNEFEQRAAGASVDAKQAVQSLTPGKPVAGTTDAN 120
QY 121 GNVTVQLPKKQNGKAVYTIKEEPKEGVAAATNMVAPFVEMIKQTDGSKYKTEBLAV 180
DB 121 GNVTVQLPKKQNGKAVYTIKEEPKEGVAAATNMVAPFVEMIKQTDGSKYKTEBLAV 180
QY 181 VHIYKPNVANDGSLHVKVGTAEENGNGAEFVLSKSESGPGTVKYIQGVXDGLYTWTT 240
DB 181 VHIYKPNVANDGSLHVKVGTAEENGNGAEFVLSKSESGPGTVKYIQGVXDGLYTWTT 240
QY 241 DKEQAKRFITGSKSYEIGENDPTEAENGTELTVKVLEVGSYTLIEEVKAPNNABELIENQTK 300
DB 241 DKEQAKRFITGSKSYEIGENDPTEAENGTELTVKVLEVGSYTLIEEVKAPNNABELIENQTK 300
QY 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKIYQISVNIPLGIADKEGD 360
DB 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKIYQISVNIPLGIADKEGD 360
QY 361 ANKYKFNVLVDKHDALTFDNTVSGEYAYALYDGVTVIAPENYQVTEQANGFTVAVNPAY 420
DB 361 ANKYKFNVLVDKHDALTFDNTVSGEYAYALYDGVTVIAPENYQVTEQANGFTVAVNPAY 420
QY 421 IPTLPFGGTLKFVYFMHLINEKADPTKGFKNENAVDNGHTDQTPPTVEVVTGSKRFIKVD 480
DB 421 IPTLPFGGTLKFVYFMHLINEKADPTKGFKNENAVDNGHTDQTPPTVEVVTGSKRFIKVD 480
QY 481 GNVTVQLPKKQNGKAVYTIKEEPKEGVAAATNMVAPFVEMIKQTDGSKYKTEBLAV 180
DB 481 GNVTVQLPKKQNGKAVYTIKEEPKEGVAAATNMVAPFVEMIKQTDGSKYKTEBLAV 180
QY 541 LKGYTYLIEETVAPDDVYLLTNRIEFVNVESGYGTENLVSPKVPNNKKGTLPSGKKG 600
DB 541 LKGYTYLIEETVAPDDVYLLTNRIEFVNVESGYGTENLVSPKVPNNKKGTLPSGKKG 600
QY 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627
DB 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627

RESULT 2
US-10-206-576-218
; Sequence 218, Application US/10206576
; Publication No. US20030017495A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06

APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D1
INFORMATION FOR SEQ ID NO: 218:
SEQUENCE CHARACTERISTICS:
LENGTH: 627 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 218:
US-10-206-576-218

Query Match 100.0%; Score 3225; DB 4; Length 627;
Best Local Similarity 100.0%; Pred. No. 9.6e-188; Indels 0; Gaps 0;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQLKQWYVSTLLILPLFTSVLGTTTAFABENGSAOLVIHKKKMTDLPDPLIONSG 60
DB 1 MKQLKQWYVSTLLILPLFTSVLGTTTAFABENGSAOLVIHKKKMTDLPDPLIONSG 60
QY 61 KEMSEFDKXQGLADYTFSIYVNTNEFEQRAAGASVDAKQAVQSLTPGKPVAGTTDAN 120
DB 61 KEMSEFDKXQGLADYTFSIYVNTNEFEQRAAGASVDAKQAVQSLTPGKPVAGTTDAN 120
QY 121 GNVTVQLPKKQNGKAVYTIKEEPKEGVAAATNMVAPFVEMIKQTDGSKYKTEBLAV 180
DB 121 GNVTVQLPKKQNGKAVYTIKEEPKEGVAAATNMVAPFVEMIKQTDGSKYKTEBLAV 180
QY 181 VHIYKPNVANDGSLHVKVGTAEENGNGAEFVLSKSESGPGTVKYIQGVXDGLYTWTT 240
DB 181 VHIYKPNVANDGSLHVKVGTAEENGNGAEFVLSKSESGPGTVKYIQGVXDGLYTWTT 240
QY 241 DKEQAKRFITGSKSYEIGENDPTEAENGTELTVKVLEVGSYTLIEEVKAPNNABELIENQTK 300
DB 241 DKEQAKRFITGSKSYEIGENDPTEAENGTELTVKVLEVGSYTLIEEVKAPNNABELIENQTK 300
QY 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKIYQISVNIPLGIADKEGD 360
DB 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVAIGEKKIYQISVNIPLGIADKEGD 360
QY 361 ANKYKFNVLVDKHDALTFDNTVSGEYAYALYDGVTVIAPENYQVTEQANGFTVAVNPAY 420
DB 361 ANKYKFNVLVDKHDALTFDNTVSGEYAYALYDGVTVIAPENYQVTEQANGFTVAVNPAY 420
QY 421 IPTLPFGGTLKFVYFMHLINEKADPTKGFKNENAVDNGHTDQTPPTVEVVTGSKRFIKVD 480
DB 421 IPTLPFGGTLKFVYFMHLINEKADPTKGFKNENAVDNGHTDQTPPTVEVVTGSKRFIKVD 480
QY 481 GNVTVQLPKKQNGKAVYTIKEEPKEGVAAATNMVAPFVEMIKQTDGSKYKTEBLAV 180
DB 481 GNVTVQLPKKQNGKAVYTIKEEPKEGVAAATNMVAPFVEMIKQTDGSKYKTEBLAV 180
QY 541 LKGYTYLIEETVAPDDVYLLTNRIEFVNVESGYGTENLVSPKVPNNKKGTLPSGKKG 600
DB 541 LKGYTYLIEETVAPDDVYLLTNRIEFVNVESGYGTENLVSPKVPNNKKGTLPSGKKG 600
QY 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627
DB 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627

RESULT 3
US-10-661-809-13
; Sequence 13, Application US/10661809
; Publication No. US20040101919A1
; GENERAL INFORMATION:
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
; TITLE OF INVENTION: POSITIVE BACTERIA

FILE REFERENCE: P07741US01/BAS
CURRENT APPLICATION NUMBER: US/10/661,809
CURRENT FILING DATE: 2003-09-15
PRIOR APPLICATION NUMBER: 60/410303
PRIOR FILING DATE: 2002-09-13
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 627
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-661-809-13

Query Match 100.0%; Score 3225; DB 4; Length 627;
Best Local Similarity 100.0%; Pred. No. 9.6e-188;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAFABENGSAQLYHKKQMTDLPDPLIONSG 60
DB 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAFABENGSAQLYHKKQMTDLPDPLIONSG 60
QY 61 KEMSEBDKXQGLADYTFISYNTNFEYBQRAAGASVDAKQAVQSLTPKQVAAQCTTDAN 120
DB 61 KEMSEBDKXQGLADYTFISYNTNFEYBQRAAGASVDAKQAVQSLTPKQVAAQCTTDAN 120
QY 121 GNVTVQLPKQNGKQAVYTIKEEPKEGVAAATNMVAFPVYEMIKQTDGSKYKGTBELAV 180
DB 121 GNVTVQLPKQNGKQAVYTIKEEPKEGVAAATNMVAFPVYEMIKQTDGSKYKGTBELAV 180
QY 181 VHIYPRNVANDGSLHVKVGTABENGGLNGAEFVLSKSESGPGTVKXIGVVDGLTYTWT 240
DB 181 VHIYPRNVANDGSLHVKVGTABENGGLNGAEFVLSKSESGPGTVKXIGVVDGLTYTWT 240
QY 241 DKEQAKRFTGKSYEIGENDPTEAENGTELTVMKLEVSYLEEVKAPNNAELINQK 300
DB 241 DKEQAKRFTGKSYEIGENDPTEAENGTELTVMKLEVSYLEEVKAPNNAELINQK 300
QY 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIEBKIKYQISVNIPLGIADKEGD 360
DB 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIEBKIKYQISVNIPLGIADKEGD 360
QY 361 ANKYVFNVLVDGDAALTFDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVAANPAY 420
DB 361 ANKYVFNVLVDGDAALTFDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVAANPAY 420
QY 421 IPTLPFGTLKRVYFMHLEKADPTKGFKNANVDNGHTDDQTPPTVEVVTGSKRFIKVD 480
DB 421 IPTLPFGTLKRVYFMHLEKADPTKGFKNANVDNGHTDDQTPPTVEVVTGSKRFIKVD 480
QY 481 GGVTAQALAGASPVVRDQNSDTANLKTDETTKATWTKABATTTTADGLVDLITG 540
DB 481 GGVTAQALAGASPVVRDQNSDTANLKTDETTKATWTKABATTTTADGLVDLITG 540
QY 541 LKRYTYYLEETVAPDDYVLLTRIRIEFVNEQSYGTENLVSPKVPNKHKGTLPTSGKG 600
DB 541 LKRYTYYLEETVAPDDYVLLTRIRIEFVNEQSYGTENLVSPKVPNKHKGTLPTSGKG 600
QY 601 IYVYLGSGAVLLIAGVYFARRRKENA 627
DB 601 IYVYLGSGAVLLIAGVYFARRRKENA 627
```

RESULT 4
US-10-912-362-218
Sequence 218, Application US/10912362
Publication No. US20050043528A1
GENERAL INFORMATION:
Applicant: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESSES:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 14200 Shady Grove Road

```
QY 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAFABENGSAQLYHKKQMTDLPDPLIONSG 60
DB 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAFABENGSAQLYHKKQMTDLPDPLIONSG 60
QY 61 KEMSEBDKXQGLADYTFISYNTNFEYBQRAAGASVDAKQAVQSLTPKQVAAQCTTDAN 120
DB 61 KEMSEBDKXQGLADYTFISYNTNFEYBQRAAGASVDAKQAVQSLTPKQVAAQCTTDAN 120
QY 121 GNVTVQLPKQNGKQAVYTIKEEPKEGVAAATNMVAFPVYEMIKQTDGSKYKGTBELAV 180
DB 121 GNVTVQLPKQNGKQAVYTIKEEPKEGVAAATNMVAFPVYEMIKQTDGSKYKGTBELAV 180
QY 181 VHIYPRNVANDGSLHVKVGTABENGGLNGAEFVLSKSESGPGTVKXIGVVDGLTYTWT 240
DB 181 VHIYPRNVANDGSLHVKVGTABENGGLNGAEFVLSKSESGPGTVKXIGVVDGLTYTWT 240
QY 241 DKEQAKRFTGKSYEIGENDPTEAENGTELTVMKLEVSYLEEVKAPNNAELINQK 300
DB 241 DKEQAKRFTGKSYEIGENDPTEAENGTELTVMKLEVSYLEEVKAPNNAELINQK 300
QY 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIEBKIKYQISVNIPLGIADKEGD 360
DB 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIEBKIKYQISVNIPLGIADKEGD 360
QY 361 ANKYVFNVLVDGDAALTFDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVAANPAY 420
DB 361 ANKYVFNVLVDGDAALTFDNTSGEYAYALYDGDVIAPENYQVTEQANGFTVAANPAY 420
QY 421 IPTLPFGTLKRVYFMHLEKADPTKGFKNANVDNGHTDDQTPPTVEVVTGSKRFIKVD 480
DB 421 IPTLPFGTLKRVYFMHLEKADPTKGFKNANVDNGHTDDQTPPTVEVVTGSKRFIKVD 480
```

CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell latitude
OPERATING SYSTEM: Windows 2000
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/912,362
FILING DATE: 06-Aug-2004
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 10/206,576
FILING DATE: 2002-07-29
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D2
INFORMATION FOR SEQ ID NO: 218:
SEQUENCE CHARACTERISTICS:
LENGTH: 627 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 218:
US-10-912-362-218

Query Match 100.0%; Score 3225; DB 5; Length 627;
Best Local Similarity 100.0%; Pred. No. 9.6e-188;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 481 GDTATQALAGASFVVRDQNSDTANYLKIDETTKATWVTKAEATFTTTADGLVDITG 540
| | | | |
DB 481 GDTATQALAGASFVVRDQNSDTANYLKIDETTKATWVTKAEATFTTTADGLVDITG 540
QY 541 LKGYTYLAEETAPDDYVLLTNRIEFVNEOSYGTENIVSPKVPNKHKGTLPSTGKG 600
| | | | |
DB 541 LKGYTYLAEETAPDDYVLLTNRIEFVNEOSYGTENIVSPKVPNKHKGTLPSTGKG 600
QY 601 IYVYLSGAVLLLIAGVFPARRKENA 627
| | | | |
DB 601 IYVYLSGAVLLLIAGVFPARRKENA 627
RESULT 5
US-09-071-035-220
; Sequence 220, Application US/09071035
; Publication No. US20020045737A1
; GENERAL INFORMATION:
; APPLICANT: Gail H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; OPERATING SYSTEM: HP Vectra 486/33
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 220:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-071-035-220
Query Match 89.7%; Score 2892; DB 3; Length 560;
Best Local Similarity 100.0%; Pred. No. 1,5e-167;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 161 FVTSKSGSPGTWVXIQQVKGDLTYWTTDKQAKRFITGSKSYEIGENDPTEAENGTELT 240
| | | | |
QY 273 VGNLEVSATYILEEVKAPNNALLENQTKPFTTEANNQTPVKTKYNDTSKYDKTTPSLD 332
| | | | |
DB 241 VGNLEVSATYILEEVKAPNNALLENQTKPFTTEANNQTPVKTKYNDTSKYDKTTPSLD 300
QY 333 GKDVAIGEKIKYQISVINIPLGIADKEGDANKYKFNLYDKDAALTFDNVTSGEAYALY 392
| | | | |
DB 301 GKDVAIGEKIKYQISVINIPLGIADKEGDANKYKFNLYDKDAALTFDNVTSGEAYALY 360
QY 393 DGDVYIAPENYQTEQANGFTVANVPAYIPILTPGGTLKFVYFMHLINEKADPTKGFQNEA 452
| | | | |
DB 361 DGDVYIAPENYQTEQANGFTVANVPAYIPILTPGGTLKFVYFMHLINEKADPTKGFQNEA 420
QY 453 NVDNQHTDDQPEPTVEVVTGGRFKNDGDTATQALAGASFVVRDQNSDTANYLKIDET 512
| | | | |
DB 421 NVDNQHTDDQPEPTVEVVTGGRFKNDGDTATQALAGASFVVRDQNSDTANYLKIDET 480
QY 513 TKAATWVTKAEATFTTTADGLVDITGKYGYLYLEETVAPDDYVLLTNRIEFVNEOS 572
| | | | |
DB 481 TKAATWVTKAEATFTTTADGLVDITGKYGYLYLEETVAPDDYVLLTNRIEFVNEOS 540
QY 573 YGTENIVSPKVPNKHKGT 592
| | | | |
DB 541 YGTENIVSPKVPNKHKGT 560

RESULT 6
US-10-206-576-220
; Sequence 220, Application US/10206576
; Publication No. US20030017495A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB369P1D1
; INFORMATION FOR SEQ ID NO: 220:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 220:

US-10-206-576-220

Query Match 89.7%; Score 2892; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.5e-167;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 33 EENGESAOQVHKKKMTDLPDLIONSCKEMSEFDCYQGLADVTFSIYVNTPEFYQORA 92
DB 1 EENGESAOQVHKKKMTDLPDLIONSCKEMSEFDCYQGLADVTFSIYVNTPEFYQORA 60
QY 93 GASVDAKAQAVOSLTPGKPAQGTDTANGNVVQLPKQNGKDAVYTIKEEPKGVAAAT 152
DB 61 GASVDAKAQAVOSLTPGKPAQGTDTANGNVVQLPKQNGKDAVYTIKEEPKGVAAAT 120
QY 153 NMVVAPEVYEMIKQTDGSKYGTGEBELAVVHIYPKVVANDGSLHVKKVGTAEENGINGAB 212
DB 121 NMVVAPEVYEMIKQTDGSKYGTGEBELAVVHIYPKVVANDGSLHVKKVGTAEENGINGAB 180
QY 213 FVYSKSGSPGVTKYIQGVKDLTYTTTDEQAKRFTGKSYEIGENDPTEANGTGEIT 272
DB 181 FVYSKSGSPGVTKYIQGVKDLTYTTTDEQAKRFTGKSYEIGENDPTEANGTGEIT 240
QY 273 VKNLEVGSYILBEVKA PNNABLIENQTKPTI EANNQTPVEKTVNDTSKYDKTTPSLD 332
DB 241 VKNLEVGSYILBEVKA PNNABLIENQTKPTI EANNQTPVEKTVNDTSKYDKTTPSLD 300
QY 333 GQDVAIGEKIKYQISVNIPLGIADKGDANKYKFNLDKHDALTFPDVNTSGEYVYALY 392
DB 301 GQDVAIGEKIKYQISVNIPLGIADKGDANKYKFNLDKHDALTFPDVNTSGEYVYALY 360
QY 393 DDDTYIAPENYQVTEQANGFTVA VNPAYI PTLTPGTLKFPVFMHLEKADPTKGRKNA 452
DB 361 DDDTYIAPENYQVTEQANGFTVA VNPAYI PTLTPGTLKFPVFMHLEKADPTKGRKNA 420
QY 453 NVNDGHTDDQTPPYEVVTVGKRFIKVDGDTATQALAGSPVVRQNSDTANYLIKIDET 512
DB 421 NVNDGHTDDQTPPYEVVTVGKRFIKVDGDTATQALAGSPVVRQNSDTANYLIKIDET 480
QY 513 TKAATWVKTKAEATFTTTADGLVDITGLKGYTYLBEFVAPDDVYLLNRIEFVNEQS 572
DB 481 TKAATWVKTKAEATFTTTADGLVDITGLKGYTYLBEFVAPDDVYLLNRIEFVNEQS 540
QY 573 YGTENLVSPKVPKHKGT 592
DB 541 YGTENLVSPKVPKHKGT 560
```

RESULT 7
US-10-912-362-220

Sequence 220, Application US/10912362
Publication No. US20050043528A1
GENERAL INFORMATION:
APPLICANT: ChOI et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 2000
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/912,362
FILING DATE: 06-Aug-2004
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 10/206,576
FILING DATE: 2002-07-29

APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D2
INFORMATION FOR SEQ ID NO: 220:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 220:
US-10-912-362-220

Query Match 89.7%; Score 2892; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.5e-167;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 33 EENGESAOQVHKKKMTDLPDLIONSCKEMSEFDCYQGLADVTFSIYVNTPEFYQORA 92
DB 1 EENGESAOQVHKKKMTDLPDLIONSCKEMSEFDCYQGLADVTFSIYVNTPEFYQORA 60
QY 93 GASVDAKAQAVOSLTPGKPAQGTDTANGNVVQLPKQNGKDAVYTIKEEPKGVAAAT 152
DB 61 GASVDAKAQAVOSLTPGKPAQGTDTANGNVVQLPKQNGKDAVYTIKEEPKGVAAAT 120
QY 153 NMVVAPEVYEMIKQTDGSKYGTGEBELAVVHIYPKVVANDGSLHVKKVGTAEENGINGAB 212
DB 121 NMVVAPEVYEMIKQTDGSKYGTGEBELAVVHIYPKVVANDGSLHVKKVGTAEENGINGAB 180
QY 213 FVYSKSGSPGVTKYIQGVKDLTYTTTDEQAKRFTGKSYEIGENDPTEANGTGEIT 272
DB 181 FVYSKSGSPGVTKYIQGVKDLTYTTTDEQAKRFTGKSYEIGENDPTEANGTGEIT 240
QY 273 VKNLEVGSYILBEVKA PNNABLIENQTKPTI EANNQTPVEKTVNDTSKYDKTTPSLD 332
DB 241 VKNLEVGSYILBEVKA PNNABLIENQTKPTI EANNQTPVEKTVNDTSKYDKTTPSLD 300
QY 333 GQDVAIGEKIKYQISVNIPLGIADKGDANKYKFNLDKHDALTFPDVNTSGEYVYALY 392
DB 301 GQDVAIGEKIKYQISVNIPLGIADKGDANKYKFNLDKHDALTFPDVNTSGEYVYALY 360
QY 393 DDDTYIAPENYQVTEQANGFTVA VNPAYI PTLTPGTLKFPVFMHLEKADPTKGRKNA 452
DB 361 DDDTYIAPENYQVTEQANGFTVA VNPAYI PTLTPGTLKFPVFMHLEKADPTKGRKNA 420
QY 453 NVNDGHTDDQTPPYEVVTVGKRFIKVDGDTATQALAGSPVVRQNSDTANYLIKIDET 512
DB 421 NVNDGHTDDQTPPYEVVTVGKRFIKVDGDTATQALAGSPVVRQNSDTANYLIKIDET 480
QY 513 TKAATWVKTKAEATFTTTADGLVDITGLKGYTYLBEFVAPDDVYLLNRIEFVNEQS 572
DB 481 TKAATWVKTKAEATFTTTADGLVDITGLKGYTYLBEFVAPDDVYLLNRIEFVNEQS 540
QY 573 YGTENLVSPKVPKHKGT 592
DB 541 YGTENLVSPKVPKHKGT 560
```

RESULT 8
US-10-661-809-19

Sequence 19, Application US/10661809
Publication No. US20040101919A1
GENERAL INFORMATION:
APPLICANT: HOOK, Magnus


```
/ TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
/ FILE REFERENCE: P07741US01/BAS
/ CURRENT APPLICATION NUMBER: US/10/661,809
/ CURRENT FILING DATE: 2003-09-15
/ PRIOR APPLICATION NUMBER: 60/410303
/ PRIOR FILING DATE: 2002-09-13
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 19
/ LENGTH: 625
/ TYPE: PRT
/ ORGANISM: Staphylococcus epidermidis
US-10-661-809-19

Query Match
Best Local Similarity 75.8%; Score 2446; DB 4; Length 625;
Matches 462; Conservative 64; Mismatches 95; Indels 2; Gaps 1;

QY 1 MKQLKKVWYVSTLLILPLFTSVLGTTTAPFAENGESQQLVIHKKKMTDLPDPLIONSG 60
DB 1 MKNHKKINVMGLVLELILPLTNSFGAKKVFABEET--AAQVILHKKKMTDLPDPLIONSG 58
QY 61 KEMSEFDKYQGLADYTFISLYNTNEFEQRAAGASVDAKQAVQSLTPGKPVAGSTTDAN 120
DB 59 KEMSEFDQYQGLADISFSYVNTTQGFYQGRDGASVDAKQAVQSLTPGTPVASGTTDAD 118
QY 121 GNVYVQLPRKQKQAVVTIKKEPKEGVVAATMNVVAPFVYEMIKQTDGSKYKGBELAV 180
DB 119 GNVVLSLPRKQKQAVVTIKKEPKQGVAAANMVLAFVYEMIKQADDSYKGBELDT 178
QY 181 VHIYPRNVVANDGSLHVKVKTGAENGLNGAEFVVISKSGSPGVYKIQGVKDGXYTWT 240
DB 179 IHLVYKRYVNDGTLKATKGTGAENGLNGAEFVVISKSGSPGVYKIQGVKDGXYTWT 238
QY 241 DKEQAKRFTGSKSYEIGENDPTEAENGTELTVKNLEVGSYLLEEVKAPNNAELLLENQTK 300
DB 239 DQTKAKHFTGHSYDIGNNDFAEASIEKGQLIYNHLEVGKYNLEEVKAPDNAMETEKQTI 298
QY 301 TPFTLEANNQTPVEKTVKNDTSKDKTTPSLDGKQVATGEKIKYQISVNIPLGIADKED 360
DB 299 TPFTLEANSQTPVEKTIKNDTSKVDKTPQLNGKQVATGEKIKYQISVNIPLGIADKED 358
QY 361 ANKYKFNVLVDGDAALTFDNTSGEYAYALYDGDVLAPENVQVTEQANGFTVAVNPAY 420
DB 359 QNKYTFPLIDHDALTFDNDSSGTYAALYDKKEIDPVVSYTEQDGFVSDVDPNY 418
QY 421 IPTLPFGTLKVFYFMHLEKADPTKGFKNKANVDNGHTDQTPPTVEVVTGAKRPIKYD 480
DB 419 IPSLTPGGLTKVFYVYHLEKADPTKGFSGNQANVDNGHTNDQTPPSVDVVTGAKRPFVKVD 478
QY 481 GDNVTAQALAGSFVVRDQNSDTANYLIKIDETTKAATWTKAKAETFTTTADGLVDING 540
DB 479 GDNVTDQTLAGAEFVVRDQSDTAKYKLSIDPSTKAVSWSAKESAVFTTTSNGLDVVG 538
QY 541 LKYGTYVEEYVAPDDYVLLTRNLEFVNVEQSGTTEENVSPKPKKGLPSTGKG 600
DB 539 LKYGTYVEEYVAPDDYVLLTRNLEFVNVEQSGTTEENVSPKPKKGLPSTGKG 598
QY 601 IYVYIGAGVALLIAGVTFARRR 623
DB 599 IYVYIGAGVALLIAGVTFARRR 621

RESULT 9
US-10-661-809-11
/ Sequence 11, Application US/10661809
/ Publication No. US20040101919A1
/ GENERAL INFORMATION:
/ APPLICANT: HOOK, Magnus
/ TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
/ FILE REFERENCE: P07741US01/BAS
```

```
/ CURRENT APPLICATION NUMBER: US/10/661,809
/ CURRENT FILING DATE: 2003-09-15
/ PRIOR APPLICATION NUMBER: 60/410303
/ PRIOR FILING DATE: 2002-09-13
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 11
/ LENGTH: 476
/ TYPE: PRT
/ ORGANISM: Staphylococcus epidermidis
US-10-661-809-11

Query Match
Best Local Similarity 13.6%; Score 438.5; DB 4; Length 476;
Matches 166; Conservative 70; Mismatches 211; Indels 193; Gaps 22;

QY 8 WYTSVTLILPLFTSVLGTTTAPFAENGESQQLVIHKKKMTDLPDPLIONSGKEMSE 65
DB 6 WLSICVMLLALFGFSQ-----QALAEASQASQVTLHKLPLPDGQLPEQO--QNTGEEGTL 59
QY 66 FDYQGLADYTFISLYNTNEFEQRAAGASVDAKQ--AVQSLTPGKPVAGSTTDA---- 119
DB 60 LQNYGLNDVITYQVYDVTDPFYQLRSEKGTQVQAEQRLAETGATNRKRPDAEDTKQTINGE 119
QY 120 NGNVTVQLPRK--QNGKDAVYTIKEBPKEGVAAATMNVVAPFVYEMIKQTDGSKYKYTE 176
DB 120 DGVVFSFLASKDSQGRDRAVLYFEVAEAPVEYKASNLVVLIPVQDPQGS----- 170
QY 177 ELAVVHIYPRNVVANDGSLHVKVKTGAENGLNGAEFVVISKSGSPGVYKIQGVKDGLY 236
DB 171 -LTHIHLVYKPN----- 180
QY 237 TWTTDKEQAKRFTGSKSYEIGENDPTEAENGTELTVKNLEVGSYLLEEVKAPNNAELLLE 296
DB 181 -----BENAYDL----- 187
QY 297 NOTKTPTEANNQTPVEKTVKNDTSKVDKTPPSLDGKQVATGEKIKYQISVNIPLGIAD 356
DB 188 -----PLEKTVLDRKQGFNQ-----GEHINYQLTQIPANILG 221
QY 357 KEGDANKYKRNVLVDKHDALTFDNTSGEYAYALYDGDVLAPENVQVTEQANGFTVAV 416
DB 222 -----YQEFRLSKADATLTTL-----LPESIEVYKAGTVYV--GYTLTQKHGFTLDF 268
QY 417 NPAYIPTLPFGTLKVFYFMHLEKADPTKGFKNKAN--VDNGHTDQTPPTVEVVTGKGR 475
DB 269 SIKDLQNA--NOTWTVSYQMLKETAEPDTAINNEGQVLTDKHT--LTKRAIVRTGKS 324
QY 476 FIKVDGVTATQALAGSFVVRDQNSDTANYLIKIDETTKAATWTKAKAETFTTTADGL 535
DB 325 FVKVDSE--NAKITLPEAVFIVNQAGEYLN-----ETANGYRQKEKALAKKFTSNQAGE 378
QY 536 VDTGLKTYGTYLEETVAPDDYVLLTRNLEFVNVEQSGTTEENVSPKPKKGLPSTGKG 592
DB 379 FSVKGLKQGYFLEISAPKGYLQTEIPFTVGNKSAVYNGQRTAPLHVINKKVESGF 438
QY 593 LPSTG--GKIYVYIGAGVALLIAGV-----YFARRRKE 625
DB 439 LPKTNERSIMWTING---LLIIGVYVIMLFYQKQAGE 474

RESULT 10
US-10-661-809-17
/ Sequence 17, Application US/10661809
/ Publication No. US20040101919A1
/ GENERAL INFORMATION:
/ APPLICANT: HOOK, Magnus
/ TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
/ FILE REFERENCE: P07741US01/BAS
/ CURRENT APPLICATION NUMBER: US/10/661,809
/ CURRENT FILING DATE: 2003-09-15
/ PRIOR APPLICATION NUMBER: 60/410303
```

PRIOR FILING DATE: 2002-09-13
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.1
SEQ ID NO 17
LENGTH: 473
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-661-809-17

Query Match 12.2%; Score 395; DB 4; Length 473;
Best Local Similarity 24.4%; Pred. No. 9, 6e-16;
Matches 156; Conservative 79; Mismatches 200; Indels 204; Gaps 26;

QY 4 LKKV-WYWTST-LILLPLFVSUGTTTAFENSGSAQLVHKKKMTD--LP--DPLQ 57
DB 1 MKKLGMLSCFLFLFKPAFTOVA-----TETETMVOITLHKLLPFGQLPKNHP--- 51
QY 58 NSGKMSSEFDKYQGLADVTFSIYNVTNEFEYQRAAGASVDAKQ--AVQSLTPGKPAQ 115
DB 52 NDGQKALLQYRGINGVTFQYVDYDSDFYHLREKGTVEBAQAEIAKKGASSGMTAA 111
QY 116 TT---DANGVTVOL-PRKQNGDAVYTIKEBPKEGV--AATMNVAPPYEMIKQTD 168
DB 112 TTTTLNNEBGLASFSLAARDQEKRDYALFIESKYPEVVKERAEENVVVLPHYGQ----- 166
QY 169 GSKYKYTEELAVVHIYPRKVVANDSLHVKYKGTAEENGLAEPVYSKSGSPGVAKYI 228
DB 167 ---NNQQLSTHLYPKN----- 180
QY 229 QGVKDLVYTWTDKQAKRFTGKSYEIGENDPTEANGTGLTVNMLEVGSYILEVYA 288
DB 181 -----ENDYPD----- 187
QY 289 PNNALLENQKTPPTTEANNQTPVEKTVKNDTSKVDKTPSLDGDVAIGEKIKYQISV 348
DB 188 -----PPEKYLEBPPN-----DPTIGEKITYSLHT 213
QY 349 NPLGLADKESGANKYKVKNLVDKDAALF--DWTSGEYAVALYDGGTVLAPENYQVT 406
DB 214 TTPVILD-----YQFELSDSADALFTLPNSLTSSNGEKITEG-----FVH 258
QY 407 EQANGFTVAVNPAYIPTLT--PGTLKFVYFMHLNEKADPTGKPKNEAVNDGHTDDQTP 464
DB 259 KKHPEFDVLF--IPSLKAVAGKGLTISYQQLSTQANKEINNNGLDGPF--GVST 313
QY 465 PTEVAVTGKRFIKVDGDTATQALAGASFVVRDQNSDTANYLIKIDETTKAATWYTKAE 524
DB 314 KKVSVYTGSKQFVKIETN--KPKRLAGAVFLIKNK--AGNYIQ--QTANGYKMTKNESD 367
QY 525 ATFTTTADGLVDIGLAKKGYTLEETVAPDDYVLLTNRIEFVNVAGSGTTEENLVSPK 584
DB 368 ALHLISDKNGAFSISGLKGTGYRLKRIEAPSGYILSETEIPTTIS--TFLSEDKADSLIK 426
QY 585 VPKKHKGT---LPSTGSKGIYYVLGSGAVLLIAGYFA 620
DB 427 VVKKENSRPFLPKTN-----ETKNTLLGVGVNFA 457
RESULT 11
US-09-071-035-274
Sequence 274, Application US/09071035
Publication No. US20020045737A1
GENERAL INFORMATION:
APPLICANT: Gill H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSES: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brooks
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 274:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-274

Query Match 10.8%; Score 347; DB 3; Length 430;
Best Local Similarity 24.0%; Pred. No. 7e-13;
Matches 142; Conservative 64; Mismatches 200; Indels 186; Gaps 19;

QY 8 WYTVSTLILLPLFVSUGTTTAFENSGSAQLVHKKKMTD--LPDPLQNSGKMS 65
DB 6 WLSICMLALFQFSQ-----QALBASQASVQVTLHKLLPFGQLPEQO-QNTGSEGT 59
QY 66 PDYQGLADVTFSIYNVTNEFEYQRAAGASVDAKQ--AVQSLTPGKPAQGTDA----- 119
DB 60 LQNTGGLADVTYQVVDYDTPFQLASBEGTVEBAQRLAETQATNRKPIAEKTYQIN 119
QY 120 NGNVTVOLPKK---ONGKDAVYTIKEBPKEGVAAATMNVAPPYEMIKQTDGSKYKTE 176
DB 120 DGVSVFSLASKDSQGRDAYLFEVBAPEVVKASNLVILPVDPQGS----- 170
QY 177 ELAVVHIYPRKVVANDSLHVKYKGTAEENGLAEPVYSKSGSPGVAKYIQGVYDGLY 236
DB 171 -LTHHLVYPKN----- 180
QY 237 TWTDKQAKRFTGKSYEIGENDPTEANGTGLTVKNLBEVGSYILEVKAAPNNAELIE 296
DB 181 -----ENAYDL----- 187
QY 297 NQKTPPTTEANNQTPVEKTVKNDTSKVDKTPSLDGDVAIGEKIKYQISVNIPLGIAD 356
DB 188 -----PPEKTVLDKQGFNQ-----GEHINQGLTTQIPANILG 221
QY 357 KEGDANKYKRVLUKDKDAALFPDVTSGEYAVALYDGGTVLAPENYQVTEDANGFTVAV 416
DB 222 -----YQFELSDKADTTLTL-----LPESIEVYAGKTVVT--GYTLTKKHGFTLDP 268
QY 417 NPAYIPTLTGGLTFVYFMHLNEKADPTGKPKNEAN--VDNGHTDDQTPTEVAVTGGK 475
DB 269 SKIDLQNR-NQTMVSYQMLREKTAEBPTAINNEGQVLTGDT--LTKRATVTRGGS 324
QY 476 PIKVDGDTATQALAGASFVVRDQNSDTANYLIKIDETTKAATWYTKAEATFTTTADGL 535
DB 325 FVKVDSR-NAKITLPEAVFIVNGAGEYLN-----ETANGYKQEKALAKKFTSNQGE 378
QY 536 VDTGLKGYTLEETVAPDDYVLLTNRIEFVNVAGSGTTEENLVSPKRVN 587
DB 379 FSVKXKXWVPLLGRN-----LCTKRLSSSESNNSFVGKFKLCNERTTN 423

RESULT 12
US-10-206-576-274
; Sequence 274, Application US/10206576
; Publication No. US20030017495A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB369P1D1
; INFORMATION FOR SEQ ID NO: 274:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 274:
US-10-206-576-274
Query Match 10.8%; Score 347; DB 4; Length 430;
Best Local Similarity 24.0%; Pred. No. 7e-13;
Matches 142; Conservative 64; Mismatches 200; Indels 186; Gaps 19;
QY 8 WTVSTLLILPLFSLVLTGTTTAPFENGENSAQLVIHKKKMTD--LPDPLINGSGEMSE 65
DB 6 WLSICVMLALPFGSQ-----QLALPASQASVQVTLHKLLPFGQLPEQQ-QNTGSEGTL 59
QY 66 PDKYQGLADVTFSIYVNTNEFYEQRAAGASVDAAKQ--AVOSLTPGKPYAOGTTDA----- 119
DB 60 LQNYGLNDVTVQVVDVTPFYQLRSEGTVCQAGQLAETGATNRKKPLAEDKTYQINGE 119
QY 120 NGNVTYQLFKK---QNGKDAVYTIKEEPKEGVVAATNMVYAFPYEMIQTDGSIYKGYE 176
DB 120 DGVVSFSLASKSQCDKAKYLLFPEAEAPPEVVEKASNLVVLFPVDPQSGS----- 170
QY 177 ELAVVYIYPRNVAVANGSLHVKKVGTAEHBGLNGAFVLSKSGSGGTWKYIQGVADGLY 236
DB 171 -LTHHLVYPRN----- 180
QY 237 TWTTDEQAQKRFITGSKSYEIGENDFTEANGTGLTVKNLVEGSIYLIIEVKAQNNALIE 296
DB 181 -----EENAYYL----- 187
QY 297 NOTKTPFTLEANNQPEVETKVDKDSKVDKTPPSLDGKQVADGEKTKYQISVNIPIGLAD 356

DB 188 -----PPEKTVLDKQGFNQ-----GEHINQLTQIPIANILG 221
QY 357 KEGDANKYKFKVLNKHDAALTFEDNVTSGEYVALYDGDVTVAIPENYQVTEANGPFTVAV 416
DB 222 -----VOEFLRSKADPTLTL---LPESIEVKVAGKTVTT-GYTLITDKHGFTLDF 268
QY 417 NPAIYPTLPGGTLKFEYFMHLINEKADPTGFKNEAN-VDNHGTDDQGPPEVVEVVGKR 475
DB 269 SIKDLQNTA-NQTMWVSQYMRLEKTAEPDTALINNEGQLVTDGTT--LTKRATVATGSKS 324
QY 476 FIKVDGDVTAITDAGASFVVRDQNSDTANVYKIDETTKAATWYTKAEATFTTAQGL 535
DB 325 FVKVDSE-NAKITLPEAVFIVKNQAGEYLN-----ETANGYRQKEKALAKKFTSNQAGE 378
QY 536 VDTGKYGTYYLEETVAAPDDVLLTNRIEFVFNESQYGTENTLVSPEKVPN 587
DB 379 FSVKXKXKMPVILGRN-----LCTKRLSSNSNNSPFGGKKPLCNERTTN 423
RESULT 13
US-10-912-362-274
; Sequence 274, Application US/10912362
; Publication No. US20050043528A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 14200 Shady Grove Road
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 2000
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/912,362
; FILING DATE: 06-Aug-2004
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 10/206,576
; FILING DATE: 2002-07-29
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB369P1D2
; INFORMATION FOR SEQ ID NO: 274:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 274:
US-10-912-362-274
Query Match 10.8%; Score 347; DB 5; Length 430;
Best Local Similarity 24.0%; Pred. No. 7e-13;
Matches 142; Conservative 64; Mismatches 200; Indels 186; Gaps 19;
QY 8 WTVSTLLILPLFSLVLTGTTTAPFENGENSAQLVIHKKKMTD--LPDPLINGSGEMSE 65

Db 6 WLSICMMLALFPGSQ-----QALAAASQASVQVTLHLKLLFPDGLPEQO-ONTGSEGL 59
Qy 66 PDYQGLADPTESYNNMNEFEQRAAGSVDAAQ--AVQSLTRPKPAQSTTDA----- 119
Db 60 LQNYGLADVTYQVVDVDPFYQLRSEGTQVQEARQLAETGATNRKPIAEDKTOTINGE 119
Qy 120 NGNVAVOLPKK---ONGKDAVYTIKEEPKEGVAAATNNVAFPVYEMIKQTGSGYKYE 176
Db 120 DGVSFSLASKDSQGRDKAYLFVEAFAPEVNVKESNLVVIIPVQDPQOS----- 170
Qy 177 ELAVVHIYPKNVANDGSLHVKKVGTAENEGANGAEFVLSKSESGPVTKYIQGVKGLY 236
Db 171 -LTHLHLYPKN----- 180
Qy 237 TWTPTDKEQAQRITGKSYEIGENDPTBAENGTELTVKLYEVSYYLBEVAPNNAELIE 296
Db 181 -----EENAYDL----- 187
Qy 297 NQTKPTTEANNQTPVEKTVKNDTSKVDKTPSLDGDQVAIGEKIKYQISVNIPLGIAD 356
Db 188 -----PPEKTVLDDKQSGFNQ-----GEHINVOITTOIPANILG 221
Qy 357 KEGANKYKFNLYDKHDAALTFDNTSGEYAYALYDGDVTIAPENVQTEQANGFTYAV 416
Db 222 -----YQEFRISDKADTTLTL-----LPESIEVKVAKTVTT--GYTLTQKHGFTLDF 268
Qy 417 NPAYIFLTPGGLTKFYVFMHNEKADPTKSGFKNEM--VDNGHTDDQTPPYEVYTGGR 475
Db 269 SIKDQONRA-NQTMVVSQMRLEKTAEPPTAIINBQALVTDKGT--LTKRAVTRGSGS 324
Qy 476 FIKVDGVTATQALAGASPVVRDQNSDPTANYLKIDETTKATWVKTAETFTTTADGL 535
Db 325 FVKVUSE-NAKTLTBEAYIVIKQAGEYLN-----ETANGYRQKEKALAKFTSNGQS 378
Qy 536 VDTGLKGTGYLLEETVAPDDYVLLTNRIEFVNVESYGTENTLVSPEKVPN 587
Db 379 FSVKXKKMPVLLGRN-----LCTKRLSSSNRNSFYGGKKFLCNERETT 423

RESULT 14
US-10-333-002-16
; Sequence 16, Application US/10333002
; Publication No. US20040071729A1
; GENERAL INFORMATION:
; APPLICANT: Adderson, Elizabeth
; APPLICANT: Bohnsack, John
; TITLE OF INVENTION: Group B Streptococcus Polypeptides Nucleic Acids and Therapeutic
; TITLE OF INVENTION: Compositions and Vaccines Thereof
; FILE REFERENCE: 2511-1-001 (SJ-0039)
; CURRENT APPLICATION NUMBER: US/10/333,002
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: PCT/US01/24795
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 09/634,341
; PRIOR FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 16
; LENGTH: 502
; TYPE: PRY
; ORGANISM: Streptococcus agalactiae
US-10-333-002-16

Query Match 10.4%; Score 334; DB 4; Length 502;
Best Local Similarity 26.3%; Pred. No. 5,3e-12;
Matches 150; Conservative 79; Mismatches 223; Indels 118; Gaps 26;

Qy 100 KOAVQSL-----TPGKPVAGCTTDANGVTVOLPKKONGKDAVYTIKEEPKEGV 148
Db 3 KKMISLVLASLAFGAUSPVTPIAFA--ETGTTIVQDTQKATYKAKVDAEIDAN 60
Qy 149 VAATNVVAFPVYEMIKQTGSGYKYEELAVVHIYPKNVANDGSLHVKKVGTAE-NRG 207

Db 61 VDSNMDGA--SYLIPQKEAEYKASTDPSNLF-----TTTNGGTYYVTKDITASANE- 112
Qy 208 LINGAFVLSKSGSGYTKYIQGVNDGLYTTWTDKEQARFITGKSYELGENDFEAEK- 266
Db 113 -----IATVA-----KSIISANTPV--STVIESND 136
Qy 267 GTGELTVKLYEVSYYLBEVKAAPNNAELIENQTKPTTEANNQTPVEKTVKNDTSKVDK 326
Db 137 GTEVAINVS--QGYTYVS--STVNNGAVIMTSVTP-----NATHEK--NTDALTWGS 184
Qy 327 TTPSLDGDQVAIGEKIKYQISVNIPLGIADKEGANKYKFNLYDKHDAALTFDNTSGE 386
Db 185 GSKTVQDKTVYSVGDVTKTITTKNAVNHGTE----KVVQYIKDTPMSASVVD--LNEGS 239
Qy 387 YAYALYDGDVTI-----APENVQTEQANGFTVANNPAYIPLTPTGG----- 428
Db 240 YEVTTTDSGNITTLTQSEKATYGRKNLLENNNFTTIT--DWAANTPTGTONGANDD 297
Qy 429 -----TLKFVYFMHNEKADPTKG-----FNGEANVDNGHTDQTPPYEVYTGGRPI 477
Db 298 FPKYGINITTYTTLGVLSGAKPGSADLPENNTINATINPNTSNDPQGVTVRQGITIK 357
Qy 478 KVDGVTATQALAGASPVVRDQNSDPTANYLKIDETTKATWVKTAETFTTTADGLVD 537
Db 358 KIDGSTKA--SLQGAIFVLKNA--TGQFLNFDNTN--VEW--GTBANATEYTTGADGIIT 410
Qy 538 ITGLKGTGYLLEETVAPDDYVLLTNRIEFVNVESYGT--ENLVSPEKVPKHGKT--LP 594
Db 411 ITGLKEGYVYLVEKAPLGVNLLDSQKVLIDGATDTTNSDNLVNPFTVEN--NKGTSLP 469
Qy 595 STGKGIVYLAGSAGVALLIAGVFPARRK 624
Db 470 STGGIGTTIFITIGALVIGAGIVVARRR 499

RESULT 15
US-09-071-035-276
; Sequence 276, Application US/09071035
; Publication No. US20020045737A1
; GENERAL INFORMATION:
; APPLICANT: Gili H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 276:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: amino acid

Db	309	SATVNGSTYVWVPBTNDV-KLDYGNPNPTESE-POEGPPANOEIKYIKDMAVADGTTITDYNV	367
Qy	364	YVK--FNLVVDKHDALTEFDNVTSGEYAVALVDGDTJLAPEN----YQVTEQANGFVAVNP	418
Db	368	AVKAI FTLQEKQDTG-ITWVNVASHBATKSPKFHTFTGLDNTKTYRVERVSGYT-----P	422
Qy	419	AYITTLTPGGTLLKFTVYFMHLNEKADPTKGFQNEANV-DNGHTDQT---PPTYEVVTGG	473
Db	423	EYV-----SFNGGVVITKNNKNSNDPPPIINSEBKVVYTG	457
Qy	474	KRFIKVNDVATQALAGASFVVRDONSDTAN-----YLKIDETTKA--	515
Db	458	RKPYKT--NQANTERLHAGTFLVKKEGKYLARKGAATAEAKAVTKATKALDDEAVKYN	515
Qy	516	-----AT-----WVTKA--EATFTTTADGLVD	537
Db	516	DLTKEKQEGEGKTALATVDQOKRAYNDAFVKANSYEMVADKADNVKLLISNAGQOFE	575
Qy	538	ITGKKGYYLAEETVAPDDVYLITNRJEFVUNEOSY--GTLENV-----SPBKYPN	587
Db	576	ITGLDKGTIYLSFEETQAPAGVATYLSGDVAFETANSYSYGATITDIAIDKGVAKDAQOYON	635
Qy	588	KHKGTLPBTGSKGIYVYISGAVILLIAGVYFARRKENA	627
Db	636	K-KVTIIPQTGIGITLFTIIGLSIMLGAVVMKKRQOSEEA	674

```

RESULT 2
US-10-531-659-17
Sequence 17, Application US/10531659
Publication No. US20060115479A1
GENERAL INFORMATION:
APPLICANT: InterCell AG
TITLE OF INVENTION: Nucleic acids coding for adhesion factors of group B streptococcus
FILE OF INVENTION: adhesion factors of group B streptococcus and further uses thereof
REFERENCE: I 10003 PCT
CURRENT APPLICATION NUMBER: US/10/531,659
NUMBER OF SEQ ID NOS: 258
SOFTWARE: PatentIn version 3.1
SEQ ID NO 17
LENGTH: 901
TYPE: PRT
ORGANISM: Streptococcus agalactiae
US-10-531-659-17

```

Query Match	7.1%;	Score 230.5;	DB 6;	Length 901;
Best Local Similarity	19.0%;	Pred. No. 5.3e-07;		
Matches 180;	Conservative 108;	Mismatches 287;	Indels 371;	Gaps 377;

```

QY 1 MKOLAKWYVYSTLLLI---LPLFYSVLGTTTAPAEENGSAOLVJHKKMTDLPDBL 55
Db 1 MRKYOK-FSKILLTSLFCLSGQIPLNTNVLGSEST--VPENGAKGLV--KXTDDQNKPLS 55
QY 56 ----- 55
Db 56 KATFVLKTTAHPESEKIEKVTAELTGEATFDNLIPEDTTLSEETAPDEGKKTNTQWQYKVE 115
QY 56 -----TONGSEKSEBPKYQOGLADVTF--SIYVNTNEPFYEPORAGASYDAKQAVQSL 106
Db 116 SNKGTITQNSDKNKSTIGQNHHELDKQPPGCIEDTRESYKLEHWKGVSPNGSESAKAV 175
QY 107 TPG-----KPAQGTTD-----ANGNTVVLPKQNGKDAYTTI 140
Db 176 NPYSSEGEHIREIPEGTLSKRISVEVDLAHNKRYKIELTVSGKTIYKPEVDKQKPLDVFVL 235
QY 141 KEE-----PKEGVAAATNMVVAFPVY----- 161
Db 236 DNSNSMNNDDGENFGRHNKAKKAABALGTAVADILIGANSDNRFALVTYGSDFDGRSYDVV 295
QY 162 EMIKQTDGSYKYCYTEELAVVAHVIPEKNVAVANDGSLVKKVGT-----AENEGI---NG 210

```

```

Db 296 KGFKEBDKTYGLQTKFTIQTENYSHKOLTNNAEIIKKIPLTEAPRAKMGSTINGLTPBEOQ 355
Qy 211 AEFVYSK-----SEGSPEGVYKIQGVNDGLYTWTDEQAKRFITGKS 253
   :::::      :::::      :::::      :::::      :::::      :::::
Db 356 KQYLSKGTGEFTWKAFAFEBADILISQVBRNSOKILVHTDGV---PTRSYALNFPGLGAS 412
Qy 254 YE-----IGENDFTBAEN-----GTGE-----LTYKNEVGSYILE 284
   ||:      :::::      :::::      :::::      :::::      :::::
Db 413 YESQFQMKQKNGYLNKSNFLLTDKPEDIKGNESYFLPPLDSYQOILISGNLOKLHYLDL 472
Qy 285 EVKAP-----NNALIEHOTKPTFTTEANNOT----- 311
Db 473 NLNAPKGTIYRNGPAREHGTPTKLYLNSLKQGNVYIPNPGIDISAPROVYNEDYKKNQDS 532
Qy 312 -----PYEKTYKNDTSKVDKTPPSLDGKDAIGE---KIKQYISV--- 348
Db 533 TFQKLKEAPFELSDGEITELMKSFSSKDEYVTPYITSSDASNNELISKQOQFEKVLTKE 592
Qy 349 -NIFLG-IADKEGDANKYKERNL-----VDKHDAAFLPDNVTSGEYAVYALY 392
   :::::      :::::      :::::      :::::      :::::      :::::
Db 593 NSIYNGTIEDPMGD-----KINLOKNGQTLQPSDYTLQSGNGSITMKOSIAG--- 640
Qy 393 DGDFTYIAPENQVTEQANGFTVAVNPATYP-----TLAFGCTLKFFYFPHLINEKADP 444
Db 641 -----GNN-----DGIILKGVLEBIIKKNLYRGALNBSGQAVTLYTDVKLDSFIS 688
Qy 445 TKGFKEANAVDNGHT-----DDQTPTV-----EVTGSKRFITY 479
   N:YD      :::::      :::::      :::::      :::::      :::::
Db 689 NKFPYD-----NGRFTLNPKSEDPNTLDPPIPKIRDVAREYPTIITIKNEKKLGEIEFTYK 743
Qy 480 DGDVTAIQAALGASVVAVDONSPTANVLIKIDETTKGAATWVTKKABATFTTTADGIVJDT 539
   ||:      :::::      :::::      :::::      :::::      :::::
Db 744 DKD--NNKLLKGAFTFELQEFNEDYKLYLP-----KNNNSKVVTGSENGKISYK 790
Qy 540 GLKTYTYLEETVAAPDDVYLLTNRIEFVFNNEOSYGTENIVSPEK-----YV 586
Db 791 DLKQSKQYLIENVSKDYKQKITNK-PIITFEVVKSSIGNILIANVKNQISYHEHGDKHLT 849
Qy 587 NKH---KQTLPSGTGKI--YVYLGSAVLLLIAGVYPAARRKXENA 627
Db 850 MTHIHPKGIIMPTGKGILSIFILG-GSMMSIAGIUYIMKRYKKS 894

```

```

RESULT 3
US-10-471-571A-3352
/ Sequence 3352, Application US/10471571A
/ Publication No. US20060115490A1
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SpA
/ TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE: P026927WO
/ CURRENT APPLICATION NUMBER: US/10/471,571A
/ CURRENT FILING DATE: 2003-09-12
/ PRIOR APPLICATION NUMBER: GB-0107661.1
/ PRIOR FILING DATE: 2001-03-27
/ NUMBER OF SEQ ID NOS: 5642
/ SOFTWARE: Seqwinn9, version 1.03
/ SEQ ID NO 3352
/ LENGTH: 1349
/ TYPE: prt
/ ORGANISM: Staphylococcus aureus
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (1)..(1349)
/ OTHER INFORMATION: Ser-Asp rich
US-10-471-571A-3352

```

Query Match 4.7%; Score 153; DB 6; Length 1349;
 Best Local Similarity 21.9%; Pred. No. 0.078;
 Matches 147; Conservative 78; Mismatches 263; Indels 182; Gaps 37

46 KQMTDLPDLLIONSKEN-SBFDKYGLADVTFSIYVTFE-----YEORAAAS 95

Db 312 KNIGDKDP---NNGETIATAKHDITANLITTYFTDY--VDRENSVOMGINSIVMDADT 366
Qy 96 VDAKQAVOSLTPGKGVAGGTPDANGNVYQLP-----KKONG 133
Db 367 IIVSKNDVFP-----NVITGNTTKTTTAIQQPDVYVNEKNSIGAFETVSHVGNKPN 421
Qy 134 KDAVYTIKEEPKEGVAAATNMV-----AFP-----YEMIK--QTDGSK 172
Db 422 GYTKQTIYVNPBENSILNAKLKQAVHSSYPNNIGQINDVDIKIYQVPGKTYTLNKGID 481
Qy 173 YGTEELA-VVHIYPRKVVANDGSLHVKVGTAE--GNGAEFVLSKSGSPGVYKYIQ 229
Db 482 VNTKELTDVINOIQLGITGDNNSAVIDPGNADSAVVMVMTQFYTSE--SPTLVQMA- 539
Qy 230 GYKQGLYTTTDEKQAKRITKGSYRIGENDFTEANGELTVKRLYEGSYLLEEVKAP 289
Db 540 -----TLSTGNKSVSTGNALGFTNNO-----SGAGQGEVYK--IGNVYMEDTNK 583
Qy 290 NNAEL-----IENQTKPTTEANNOTPV-EKTVNDTSKYDKTTPSLDGDVAIGKIKY 344
Db 584 GVOELDEKGVAGVTYVF--DNTNTKVGBAVTKEDGSYLIPLPNGD-----Y 630
Qy 345 QISV-NIPLG---IADKEGDANKYVKNLVKIDALTFPDNYS--GEYAYALYDGTVI 398
Db 631 RVFENSLPKGYEVTPSKQGN-NBELDSNGLSVITVNGKONLSADLGIVKPKYNLGDYVW 689
Qy 399 ABEVY-----QVTEQANGTVAV--NPVYIPLLTGSGTLKFYFMHLNEKADPTKG-FK 449
Db 690 EDTNNGIODDEKIGSGVTATLKONGVNLKTVTTDADGKYF-----TDLNNGNYK 742
Qy 450 NEAVNNGHTDDQTPTEVTVGSGKRFIVDGDVTAQLAGA-----492
Db 743 VEPPTPEGTT-----PT--TVTSGSDIEKDSNGLITTVGIVNGDNNTLDSGFKTPKYNL 795
Qy 493 -SPVVRDONSPTANYLIKIDETTKAATWVTKAE-----ATFTTADGLVDITGLKGYT 545
Db 796 GNVYMEDTKDG--KQDSTEGISGVTVTLKNGEVLQITKTDKQKGYQFTGLENGT 851
Qy 546 YLLE-ETVAPDDVYLITNRIBFVVAEQSITTEVLVS-----PEKVPNKGKGLPS- 595
Db 852 YKVEEET-PSGY-----TPQVGSGETDEGIDSNSTGTVGKDKDNIDTISGFYK 900
Qy 596 -TGKGIYVY 604
Db 901 PTVNLGDYVW 910

RESULT 4
US-10-471-571A-2278

; Sequence 2278, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927MO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqMin99, version 1.03
; SEQ ID NO 2278
; LENGTH: 2478
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1) - (2478)
; OTHER INFORMATION: FmbB protein [Staphylococcus
US-10-471-571A-2278

Query Match 4.6%; Score 149.5; DB 6; Length 2478;
Best Local Similarity 21.7%; Pred. No. 0.32;

Matches 136; Conservative 79; Mismatches 237; Indels 175; Gaps 34;
Qy 58 NSGKENSEPDKVQGLADVTSIYNTNEFEQRAAGS-VDAKQAVOSL--TPGKPVQ 114
Db 1676 NAVKTFSEYK--DALAKIEDANNAKNEADNSASTSSIEAKQGLAEKQIADQVNO 1734
Qy 115 GTT-----DANGNVYQLPKONGKDAVTTIEEPKEGVAAATNMVAVPVE 162
Db 1735 AISKQDIEVQHNDNDINIDYITIPGKESATITDLAYAVDQKNNISLSTN-ATQEKQ 1793
Qy 163 MIKQTDGSKYGTSEL-----AVVH-----IYKRVVANDGSLHYK 198
Db 1794 AIKQVDQVQVTALESINNGVDNDVDALDQGLAIDAQVATVPK--ANQ-AIEYK 1849
Qy 199 KYGTAEENGLNABEPISSSESPGVYKI-----QGVQGLYTTTDEKQAKRITGKS 253
Db 1850 AEDTESISDQSLTAEBETE-ALMIKQITQAKQGITDA--TTTAEVEKAK---AAG 1902
Qy 254 YEIGEN---DTEAENGTEL--TVNLEVGSYI-----LEEVKAPNNA--ELIENQTK 301
Db 1903 LEAPNDIQDSTBKQALIELELTALDQIAGVAVNADATTEKEKFTNLEDLISKATE- 1961
Qy 302 PTTIANNQTPVEK--TVKN--DTSKYDKTTPSLDGKV-AIGEKIKYQISVNIPLGIA 355
Db 1962 ----DISQTTNABEATVNGSALBQKQRIINPEYKNNLBALIREVVKQIEI--IKNA 2014
Qy 356 DKEGDANKYVKNL-----VDKIDALTFPDNYSGEVAYALYDGDVIAEBNYVTEQ 408
Db 2015 DADSAKEIAEATDLGRYFDRFADKLDQOT-----NAEVAEL 2051
Qy 409 ANGFTVAVPAVYIPLLTPGSKLFYFMHLNEKADPTKFKNE--ANYDNGHTDDQ-- 463
Db 2052 QN-VTIPALVLPQNDP-----DANTNNGIDNNDAFANNAATPEPTQ 2097
Qy 464 PTVVEVYVGGKRFK-----VDGDTAQAALAGSPVVRDONSPTANYLIKIDETK 514
Db 2098 PNVSEFTANGKADASPTTPNNSDAATGETTATSATDADANKQANNSS-----VDASTN 2152
Qy 515 AATW-----VTKAEATTTTADGLVDITGLKGYIIEVYAPDDVYLITNRIBFVN 569
Db 2153 SPTMNDVTSKEVSEYNNGT-----DKPVTETDN--ATPAESTTN 2192
Qy 570 EOSYGTTEVLVSPEKVPNKGKGLPST 596
Db 2193 NNSTTATN---ENAPGSTATAPTT 2215

RESULT 5
US-10-471-571A-860

; Sequence 860, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927MO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqMin99, version 1.03
; SEQ ID NO 860
; LENGTH: 2186
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1) - (2186)
; OTHER INFORMATION: Mtp protein [Staphylococcus
US-10-471-571A-860

Query Match 4.5%; Score 146; DB 6; Length 2186;
Best Local Similarity 20.3%; Pred. No. 0.44;

	Matches	140;	Conservative	89;	Mismatches	269;	Indels	192;	Gaps	31.
Oy	23	SVLGTTTAPAEENGESAOVLVHKKQMTDLPDLIONSGEKSEMFDPYOGADVTFSIYN	:	:	:	:	:	:	:	82
Dd	1599	SIKDTPHATVDLDELDAENOLISPTLKQAQOE--INENTODAVALTD-----VANQ	:	:	:	:	:	:	:	1644
Oy	83	TNEFEQ-----RAAGASV-----DAKAQAVSLPFGKVAGCT-----	:	:	:	:	:	:	:	116
Dd	1645	TTKALEQLPKVKRRRGAALDISIEENKKOGLDIRTLDTODDERVALDITLANKIVNTKN	:	:	:	:	:	:	:	1704
Oy	117	-----TDANGNVTVQLPKKONGKQCAVYTIIKEPEKGVAATNMVAFPVY	:	:	:	:	:	:	:	161
Dd	1705	DIAQKTNAAEYDRTEETDGNDINKVILPKVQ-----VKPAARQSV-----GVKAEQN	:	:	:	:	:	:	:	1751
Oy	162	EMIKOTDSYKKGCEELAVNHIYPKNV-VANDGSLHVKVGKGAHEBGLNGAEPVISSEG	:	:	:	:	:	:	:	220
Dd	1752	ALIDSDSSTE-BERLLAKHLVEBALWQALDOIHNARKTAQVNDNSTN-QONITSK---	:	:	:	:	:	:	:	1805
Oy	221	SPGTVKVYIQGVADGLYTWTTDEKQAKRFITGSKYEI-GENDFTEAENGTELVELVKNLEVG	:	:	:	:	:	:	:	279
Dd	1806	-----IKPATVVATAALQQIQONIAIKINKINIKANNEATDBE-----QNIAMIA	:	:	:	:	:	:	:	1844
Oy	280	SYLIEFVQAAPPNNALLENQKTPEFTI-EANNQTFVEKTVKNDISKVDKTPSLDGQOVA	:	:	:	:	:	:	:	337
Dd	1848	QVEKELIRAKQOIASAVTNADVAILHDKEKREIREIEVINRKASABEQLTFLFNDKQOA	:	:	:	:	:	:	:	1907
Oy	338	IGEKITKYQIS-----VNIPLGIDAKE---GDANKVVKFNLVDDKDPAALTFPNV	:	:	:	:	:	:	:	382
Dd	1908	IEANIQAVEBERNSILAQLQNIYPIAIGIIDDSNAQVDKTSANLQTTHLDV----	:	:	:	:	:	:	:	1962
Oy	383	TSGEYAVALYDGDVIAPENVQVTEQOANGFTVAANPAVIPTLTPGGTLKFVYMHLNEKA	:	:	:	:	:	:	:	442
Dd	1963	----HPICKPPAEKTIINDLARVITALVONYRKVS-----RMKA	:	:	:	:	:	:	:	1997
Oy	443	DPTEGFKN-EANVNDGHRTDDQPPFVENVYTGKREIKXIDGPYATQALAGASFVRDONS	:	:	:	:	:	:	:	501
Dd	1998	DALKAITPLKLOMDEBELKTARTNADVDVL-KRNVALSIDIEAV-----ITEKENS	:	:	:	:	:	:	:	2047
Oy	502	DTANYLIKIDETTKATWVTKAEATTFTTTADGLVIDTGL-KY---STYYILEETVAPDD	:	:	:	:	:	:	:	556
Dd	2048	----LLRIDNINAQ-QTYAKKFAIA-----TEQLAKVALLIDQYVADGNRMIMDEBATLND	:	:	:	:	:	:	:	2097
Oy	557	YVLTNRLEPFVNE---OSYGTENLVSPPEKV-----PRKAKGT-----LPST	:	:	:	:	:	:	:	596
Dd	2098	--IKOHQTFVDELIATLKPAAETKVSPEKEOPAPKVCPTPKKETHSRKVEKELEPNT	:	:	:	:	:	:	:	2154
Oy	597	GKGGIYVYLGSAGVALLIAGVFARRREN	:	:	:	:	:	:	:	626
Dd	2155	GSEGMDLPKPEPA-LITGALLIARRRIKN	:	:	:	:	:	:	:	2182
RESULT 6										
US-10-471-571A-4496										
Sequence 4496, Application US/10471571A										
Publication No. US20060115490A1										
GENERAL INFORMATION:										
APPLICANT: CHIRON SpA										
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS										
FILE REFERENCE: P0269237WO										
CURRENT APPLICATION NUMBER: US/10/471,571A										
CURRENT FILING DATE: 2003-09-12										
PRIOR APPLICATION NUMBER: GB-0107661.1										
PRIOR FILING DATE: 2001-03-27										
NUMBER OF SEQ ID NOS: 5642										
SOFTWARE: Seqwin99, version 1.03										
SEQ ID NO 4496										
LENGTH: 9535										
TYPE: PR1										

```

US-10-471-571A-4496
Query Match 4.5%; Score 145.5; DB 6; Length 9535;
Best Local Similarity 20.3%; Pred. No. 4.2;
Matches 140; Conservative 83; Mismatches 295; Indels 173; Gaps 30;

QY 24 VLGTTLTAFAENSGSAQLVTHKKKMTDLPRLPLQNSGKMSBFDKX---QGADLVFESLY 80
DB 1436 VKSNSTNNADDAVYKADLYLTPYGPKQYVHEHLQNTGNTTDAINIYFVPSDLVNPISIG 1495
QY 81 NVYNE--FYEORAGASVDAKAQVOSLT-PGKPAVAGTTDAN-GNVTVOLPKQNGKDA 136
DB 1496 NYTHQVFSEGETFTNTITANDNFQVSVTFVNPSTSQTGLVDNHHGVSAATAPVTSATXK 1555
QY 137 VYTIKEPKGVAAATMMVAPFVYEMIKQTDOSYKKGTEBELAVHHIYKQNVANDGSJH 196
DB 1556 TINLATDTSGNTATTSFNVY---VXPLRDKYRVGTSSTANPVRJAN-ISNNATVS 1608
QY 197 VKKGTAEENGNGCAEFVSKS-----EGSPGVKYIOGVKDXLT 237
DB 1609 QADDTIIN-SLFTFTEVPRRSYARASANEITSKTVSNVSRGNMNAVYVYTYDGTTS 1667
QY 238 WTT--DKEQAKRFTGKSYSIEGENDTEAENGTEGELTVKLEVGSYILEEV-----KAP 289
DB 1668 TVTVPVGHVPEIYVASHSHYVQGDV-PAGNGSSASADYFKLSNGSDIADATITWVSGQAF 1726
QY 290 N--NAELIENQTKPTFTEANNQTPVEKTVKNDTSKVDKTPPSLDGKDVALIGEKKYQIS 347
DB 1727 NKDNTKIGEDITVTAHLIDGETTFITKTA--TYKVATVP-----KAVFEFA 1772
QY 348 VNIPL-GIADKEGDANKYKFNLVDKHDAALTFDNTSGEYALVYDGTVJAPENNYVT 406
DB 1773 RGVLYPQVSD-----MYDAQYKYPVNNSSVS 1798
QY 407 EQANGFVAANPAVYIPTLP---GGTIKYYFMHLNE-----KADPTKGFQX 450
DB 1799 TNAQ---HNNFQVFGYGGFNKDVGVISTRLKIVTYDNGQTEDLTILSKYKPDPRIDAN 1854
QY 451 EANDNGHTD-----DQTPVEVVTGGRKRFKV-----DGDV 483
DB 1855 SVTYKAKLTNOELKNNVNLNNSSVKLFKADNPPLANTNTIHTSGGFSVTVSDALPNNGI 1914
QY 484 TATDALLA-GASFVRKDQNSDTANYLK--IDETGATVYKTKAEATFTTTADGLVDI 538
DB 1915 KAKSSIMNNVYTYDDEHQVVTYTRNSVDSNDSATYVYVPOLOATT-----EGAVPI 1969
QY 539 ---TGLKXGTYVEE-----TYA-----PPDY---VLTNRIEFPVNEQSYGTEML 579
DB 1970 KGGEGPFG--HYERFIQNPFGATYTAHMSDPYTKQVGNHKAIVATLPLNGQG-TRNV 2026
QY 580 VSPKVPNKHKGLPSTGKGIYVYLSGAV 610
DB 2027 EVPVKVPVANAAPSRHDVKGQMLTNGTDAM 2057

RESULT 7
US-10-530-879-4
; Sequence 4, Application US/10530879
; Publication No. US20060140980A1
; GENERAL INFORMATION:
; APPLICANT: GUSSE, Bengt et al.
; TITLE OF INVENTION: IMMUNIZATION OF NON-HUMAN MAMMALS AGAINST STREPTOCOCCUS EQUI
; FILE REFERENCE: 0825-0173PUS2
; CURRENT APPLICATION NUMBER: US/10/530, 879
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: PCT/SE2003/001587
; PRIOR FILING DATE: 2003-10-10
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Streptococcus equi

```

US-10-530-879-4

Query Match 4.3%; Score 138.5; DB 6; Length 657;
 Best Local Similarity 18.2%; Pred. No. 0.22;
 Matches 140; Conservative 103; Mismatches 268; Indels 257; Gaps 36;

```

QY 1 MKQLKRVYVSTLLILPLFTSVLGTTFAAENEGESQVLVHKKKMTDLPPLIONSQ 60
DB 1 LKQLTIVSV--LLVFTLSAS-----LHKRATNLSNITSLSV 39
QY 61 KEMSEFDKQGLADVTFSYVNTNEFYEQPAGASVDAKQAVOSLTPGKPVAGTTDAN 120
DB 40 ASSSLDGETTYKVAF-----DDKKQKIKAK--GDTI--EVTWPTS 76
QY 121 GNTVQ-----LPKQNGDAVYTIKEPKGVAAATNMVAPVYEMIKQTDGSKYKGT 175
DB 77 GNYIIGFKNKTIPLNIRGVN-----GTLEVTLDAVFTPNQNIETMHDVSGWG 126
QY 176 BELAVNHIYKONVANDSLHVKVGTAEENGLNGAEFVLSKSGSPGVKYIQVVDGL 235
DB 127 FDIIVANN--TQTTAETSQTTVYK-----VGNRTATIVTKPEAGTGTSSFYKTKGD-- 176
QY 236 YVMTTDEKQAKP--ITKSYEIGENDFTEAENGTELTVK---NLEVGSYIIEVYKAPN 290
DB 177 -MQPNDTEHVRWELLINNKEWVANVTVEDDIQGGQTLDMSSFDITVSGRNERFVGEN 235
QY 291 NAELEIENQKTPPTIEANNQTPVEKTVKQDTSKVDKTTPSL-----DGKDVAIIGEK 341
DB 236 --ALTEFHTFPNSVITATDNH--SVRLDQYDASQNTVNIAYKTKITDFOQEFANNSK 291
QY 342 IKYQI-----SVNIPLGIADKEGDANKYKFNLV---DKHDAALTFPN 381
DB 292 IYQILYKQVSGESNHQVANINANGV-----DSRRTSFTVKKIMDKENQ---DG 342
QY 382 VTSGEYVALYDGTVIAPENVQTE-----QANGFTVAVNPAYIPLT 425
DB 343 KRPKITVQVANDQKNDKTEIELSDTNWQASFGKLDKYDSQNKITTSVKEWVPV-- 400
QY 426 PGKITLKVYFMHLEKADPTKGFKEANEANDNG---HTDDQTPPYEVTVG----- 472
DB 401 -----GVQSQVEGDSGSGFTITVYTPREVIS--ITGQKTWMDREN 438
QY 473 --GKR---FIKVDGVTATQALAGASFVRDONSPTANYLKID--ETTKAATWVTKAB 524
DB 439 QGKPKKELTVRLANDATDKAYTAS---EQTGMKTYFTNLPRKQDQKQITTYTQEBP 494
QY 525 ATFTTTADGL-----VDITGLKY-----GTYULEETV 552
DB 495 VADYTTIQSPDITNHEVALTSLKVIKYWNKDODYYHKKRKEITILLKADGVIREHQM 554
QY 553 ABD-----DYVL-----LTNRIEFVNVEQSYG-----TTENLVSPK 584
DB 555 TDDQQKMWETPDQLPVYQTKKISYSIEBKQVAGYQAPVYEVDEGLKQVTVNTINLPSV 614
QY 585 VENKHKGTLPSTGSKGI--YVYLSG---AVLLLLAGVYFARRKEN 626
DB 615 -----KLPTGGQGVKMYLLIGGFIYAILVLIS--LYQKHAKHN 653

```

RESULT 8
 US-10-471-571A-3554
 ; Sequence 3554, Application US/10471571A
 ; Publication No. US20060115490A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SPA
 ; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE: P026927WO
 ; CURRENT APPLICATION NUMBER: US/10/471,571A
 ; CURRENT FILING DATE: 2003-09-12
 ; PRIOR APPLICATION NUMBER: GB-0107661.1
 ; PRIOR FILING DATE: 2001-03-27
 ; NUMBER OF SEQ IDS: 5642
 ; SOFTWARE: SeqMin9, version 1.03

; SEQ ID NO 3554
 ; LENGTH: 1018
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 ; FEATURE:
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (1)..(1018)
 ; OTHER INFORMATION: fibrinectin-binding protein
 US-10-471-571A-3554

Query Match 4.3%; Score 137.5; DB 6; Length 1018;
 Best Local Similarity 20.3%; Pred. No. 0.49;
 Matches 151; Conservative 87; Mismatches 276; Indels 229; Gaps 37;

```

QY 27 TTTFAEENGESQVLVHKKKMTDLPPLIONSQKEMSEFDKQGL-ADVTFSYVNTNE 85
DB 42 YTT--VEENGNSA-----TDNKTSTQTTANVNIIEEQSYNAVATVEQPSNAVTV 90
QY 86 FYE-----QRAAGASVDAKQAVOSLTPGKPVAGTTDANGVTVQ-----LPKIQ 131
DB 91 TTBAPKAVQAPQTAQAPANIETVKEEYVK--EAKRPQVKEKTTQSDNSGQROVDLPKKA 149
QY 132 NGKDAVYTIKE--EPKEGVAAATNMVAPVYEMIKQTDGSKYKGTTELAVV----- 181
DB 150 TQNVAVETQVAVQAPRTASESKPRVTRSDVAEAKASNAKAYETGDTVSKYVEIGSIE 209
QY 182 -HIYPKONVANDSLHVKVGTAEENGLNGA---EVLISKSGSPG--TVKXIQGVKDG- 234
DB 210 GHNNTKVBPVHAGQRAVLKYLKLFENGHLQGDYFDDTLSSNNVNTGVSTARKVPEIKNS 269
QY 235 -----LYTWTTDKE-----QAKRPIFGKSYEIGENDFTEA---ENG 267
DB 270 VMANGEVLEGGKIRTFPTNDIEDKVDYAELEINLFIQPKVYQNTNGQNTISTLNEBQT 329
QY 268 TGBELTVKNLE-VGSYI-----LEEYKAPNN---AELEIENQKTP----- 302
DB 330 SKELDVKYDQIGNYVANLNGSIEFPNKNANRFSHVAFIKPNNGKTTSTVVTGLTKGNS 389
QY 303 -----FTIEANNQTPVEKTVKQ--DTSKVDKTPSLQKQVVAIGEKIKYQISVNI 350
DB 390 QNGNQPKVALIFEYLGNNEDIASVYANTDTSKFEKVTSMMSG----- 432
QY 351 PLGIADKEGDANKYKFNLDKHAALTFPDNTSGEYVALY--YDGPVVIAPENVQTEQA 409
DB 433 -----NLNIQNGSSISLNIENL---DKTYVHYDSEYIANGTDEVDERTQM 474
QY 410 NGFTVAVNPAYIP--TLT--PGKITLKVYFMHLEKADP--TRGFKNKANVD-----NGH 458
DB 475 VGHPEQLYKYVDRTGLTFWDMGLVLYSNKANGNEKNGPILIONNKPEYKEDTIKETLTCQ 534
QY 459 TDDQTPPYVE-----VVTGGRKRIKVDGV-----TATQA 488
DB 535 YKKNLVTVYEEBYSSTLIDYHTAIDGGGV--VDGYLETTEBTSADIDYHTAIVDS 592
QY 489 LAG--ASFVRDONSPTANYLKIDETTKAATWVTKAEATTF-----TTTADGLV 536
DB 593 EKHGHHGYT---ESSEBSNPIDPEBETHENS--KHADVVEYBEDNPPGQVTTTSNLV 647
QY 537 D-----ITGLKXGYLYLEETVAPDDVYLLTNRIEFV-----VNEQSYGTTENLVSP 582
DB 648 EFDEBSTKGIYGAVSDHTTVEDT--KEYTTESNLELVDELPEBHGAQAPVEBITK- 703
QY 583 EKVPKH--KGTLPSTGSKGIY 602
DB 704 ----NNHHISHSGLGTEHNGHNY 722

```

RESULT 9
 US-10-505-928-449
 ; Sequence 449, Application US/10505928
 ; Publication No. US20060088532A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ludwig Institute for Cancer Research et al.

```

; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 449
; LENGTH: 3396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-449
```

Query Match 4.2%; Score 135; DB 6; Length 3396;

Best Local Similarity 18.6%; Pred. No. 4.2; Indels 189; Gaps 27;

Matches 113; Conservative 82; Mismatches 223; Indels 189; Gaps 27;

```

QY 38 SAQLVHHKKMTDLDPPLIONSCKMSEFDKYQGLADVTFSIYNVTNEFEQRAAGASVD 97
DB 2486 SYMLPLHBEQNKSSPD-----TSTLSNTVS--YERSTDSGSPD 2522
QY 98 AAKQAVQSLTPGKPYAQGTTDANGVTYOLPKKQNGDAVYTIKE---EPKEGVAAATN 153
DB 2523 RFRFEDSTL--KPRKKPTE---NIIDLDKED--KDLITITESTTILEILPELTSDKN 2575
QY 154 MVVAF-----PYVEMI--KOTDSYKKGTEBELAVNHVYPRNVVANGDSLHVKKV--GTARE 206
DB 2576 TIIDIDTKPEVEDLGMQTDIDTEPSEP-----HSDNDSNDSTOVQOYIYENAVNL 2629
QY 207 GINGAFVYLSKESGPGTV--KTIQGVKGLYTTTDDKEQAKRFITGKSYEIGENDFTEAE 265
DB 2630 SLTEETF-----EGSADVLASTQATHDSEMTY--EDRSQLD----- 2664
QY 266 NGTGLTVKNLEVGSYILEEVAPNNAE---LIENQYTPPTTEANOTPEVKTAVKNDT 321
DB 2665 -----HMGHFTTGTIPAPSTETELDVLLPRTATSLPIPKSATVIPEIIGIKAEA 2713
QY 332 SKVD--KTPSLDQKDVAGIKI-----KYQ-----ISVNIPLGIADKE 358
DB 2714 KALDMFESTTSDQALADQSEIITPLQGFERTQEEYDKKAGSPFPPSSGAEEAL 2773
QY 359 GDANKYVFNLYDKHDAALT-----FDNVTSGEYAVLYDGDVTYIABENYOV 405
DB 2774 VDHTEYLSIATTHLMDQSTEVPRDWEGSNPRYTDTTLAVSTPAKLSQGTSSPLTIYS 2833
QY 406 TEQANGFTVAANPAYIP-----TLPQCGTAKFYV-----FMELNEKADPT 445
DB 2834 GSEAGSGHEIPQPSALPGIDVGSVWSPDSFKEIHVNIEATPKPSSEBYLHITPPSLIS 2893
QY 446 KGFKAENVVDNGH---TDQTPPTVEVVTGKRFI---KYDGDVTATQALAGASFVY 496
DB 2894 PDTKLEPSDDCKPELLEEMENASPTIELINVESTTEIIOFPQNKTDQVSG----- 2942
QY 497 RDQNSDTANYLIKIDETTKAATWVTKAEATFTTTADGLVDITGLKY-----GTYYL 548
DB 2943 -----EAIKMFPIIKT--PEAGVITTTADE--IELEGATQWPHSTASATYCV 2986
QY 549 EETVAP 554
DB 2987 EAGVVP 2992
```

RESULT 10

```

US-11-256-173-28
; Sequence 28, Application US/11256173
; Publication No. US20060140979A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; TITLE OF INVENTION: Antigenic Peptides
; FILE REFERENCE: coxin
; CURRENT APPLICATION NUMBER: US/11/256,173
; CURRENT FILING DATE: 2005-10-24
```

```

; PRIOR APPLICATION NUMBER: US/10/311,879
; PRIOR FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 2659
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-11-256-173-28
```

Query Match 4.2%; Score 134.5; DB 7; Length 2659;

Best Local Similarity 19.1%; Pred. No. 3.2; Indels 155; Gaps 24;

Matches 113; Conservative 73; Mismatches 250; Indels 155; Gaps 24;

```

QY 81 NVTNEFEQRAAGASVDAAKQAV---QSLTPGKPYAQGT-----TDANGVTYOLPEK 130
DB 1551 NPTMDVNTVNOKAAVSKSTKDLDDGQQLQRAKTAHTATHTASDLNOAQKALTOQVNS 1650
QY 131 QNGKDAVYTIKEPKEGVAAATNNVAFPYVEMIKQTD-----GSYKYGTE 176
DB 1651 AQNVQAVNDIKQTTQSLMTATMTGLKRGVANHQVQSDNYVNAADTNKKNDYNNAYNHAND 1710
QY 177 ELAVVHIIP-----KNVANDGSL-----HYKKVGTARENG 207
DB 1711 IINGNAQHPVITPPSDVNNALSNVTSKEHALNGEAKLNAKQEAATALGHLNNLNAQRON 1770
QY 208 L-----NGAEFVYSKESGPGTVKY---IQGVKGLYTTTDDKEQAKRFITGKSYEIGEN 259
DB 1771 LOSQYNGAHQI---DAVNTIKQNAATNLNSMGHLROAVADKQVYK-----TE 1815
QY 260 DTEAENGTEGLTVKNLEVGSYILEEVAPNNAEIENQYTPPTTEANOTPEVKTAVK 319
DB 1816 DYADAD-----TAKQNAVNS-----AVSAETIINQTTNP--TMSYDD-----VNR 1854
QY 320 DTSKVDKTPSLDQKDVAGIKIKYQISVN--IP-----LGIAKEGDANKYVFNLYD 371
DB 1855 AHSATYSKRNALNGYEKLAQSTDAARAIIDALPHLNNQKADVKSKINAAISVAGVTYK 1914
QY 372 KHDALTFDNVTSGEYAVLYDGDVTYIABENYQ--VTEQANGFTVAANPAYIPTLPQGT 429
DB 1915 QQGTPL---NTPMGLQGAINDQRTIANSQNYQDAPTPSKTATYTNVQAA----- 1961
QY 430 LKFVYFMELNEKADPTKGFKAENVVDNGHDDQTPPTVEVVTGKRFIKVDG-----D 482
DB 1962 -----KQILNKSNGQN--KTKQVTEAMQVNSAKN--NLDGTRLLDQAK 2002
QY 483 VTAATQALAGASFTVADQNSDTANYLIKIDETTKAATWVTKAEATFTTTADGLVDITGLK 542
DB 2003 QTAQOQLNNMTHLTTAQKTNLNLNQINSCTTAVAGVOTVQSN-----NTLDQAMNT--LR 2054
QY 543 YGTYYLSEETVAPDDYVLLNRIEFVNVNEQSYGTENLVSPKVPNKHKGTL 593
DB 2055 QSIANKDATKASQEDYVDANNQKQAYN--NAVAAAEITIIINANSNPEMNSTI 2104
```

RESULT 11

```

US-10-449-902-50884
; Sequence 50884, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-4020571-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
```


US-11-105-233-158

```
Query Match      3.9%; Score 124.5; DB 7; Length 2355;
Beet Local Similarity 19.9%; Pred. No. 11;
Matches 130; Conservative 82; Mismatches 245; Indels 197; Gaps 30;

QY 29 TAPA-EEENGESAKQVTHKKKQTDLPDPLIONSKESESEFQGLADYFSSYNTNFEY 87
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 879 TIYANEENQESTPPIVI-QQETGTRSDTPVSPRL-----QFVEVTDKVTIMTRP--- 930
QY 88 EORAAASVDAK-----QAVOSLRPG-----KPV-112
DB 931 ESAVGYRVDPVLPVNLPGHGORLPISRTFAEVLGLSGVYTYKVPVASHGRESKPLT 990
QY 113 AAGTDDANGNTVQVLPKQNGKDAVYTIKEPKEGVATAINNVAFPVYEMIKOD---G 169
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 991 AAGTCKLAPLNLQF---VNETDSTVLVAKMTPPRAQITGYRLTGLTRRGQPRQNVGBS 1047
QY 170 SYKYGTEELAVVHIIPKRVNVANDGSLHVKVGTAEENGINGAEFYISKEGSPGTVKYIQ 229
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1048 VSKYPLRLNLPASEETVSLVALKNGQESKATGVFTTLQPGSSIPRYNTEVETETI---- 1103
QY 230 GVKDGLVTWTDKEQAKR--FITGKSYEIGENDFTEAENGTELTVKMLEVG-STYL--- 283
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1104 ---VITWT---PAPRIGFYLGYRPSQGEAPREVTSDSGSIVVSGLTPGEVEYVYTIQ 1154
QY 284 ---EEVKAPNNALIEHQTKPT-----IEANNQTPVKEKTVKNDTSKYDKTTPSLD 332
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1155 VLRDQGEADAP---IYKVVTPPLSPPTNLHLEANPDTGV-LTVSWES---TTPDIT 1204
QY 333 GKDVAIIGEIKIKYQISVNIPLGIADKEGDANKYKFNLYDKHDAALFDNVTSG-EYAAVL 391
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1205 G-----YRITTPING---QQGNSLEEV---VHADQSSCTFDLISGLENNVSV 1247
QY 392 YDGDVIAIPENYQVTEQANGFTVAVNPAIYPTLTREGTLKFTYFMHLEKADPTGFKNE 451
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1248 Y-----TVKDKESVSPISDTIIPAVPPTDLRF-----TN 1277
QY 452 ANVDNGHTDDQTPPVEVNVTVGSKRFKY-----DGDVTAQALAGASFYVR 497
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1278 IGPDMRYTWAPPSIDLTNPLVRSFVNNEEDVALSISPSDNAVLTNLLPGEYVVS 1337
QY 498 ---DONSDT-----ANYLKIDETTKATWKTAAEATTF----- 528
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1338 VSSVVEQHESTPLRGQKTGLDPSGIDPSDITANSFTVHMLAPATIGYIRHHPHF 1397
QY 529 -----TTTADGLVDITGLKGYTYLEETVAPDDYVLLTNRIE--FVNEQS 572
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 1398 SGRPREDRVPHSRNSITLTNLTPTGEYVVSIVA-----LNGREBSPLLIGQOS 1445
```

RESULT 14
US-10-471-571A-5118

```
/ Sequence 5118, Application US/10471571A
/ Publication No. US20060115490A1
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SPA
/ TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE: P026927WO
/ CURRENT APPLICATION NUMBER: US/10/471,571A
/ CURRENT FILING DATE: 2003-09-12
/ PRIOR APPLICATION NUMBER: GB-0107661.1
/ PRIOR FILING DATE: 2001-03-27
/ NUMBER OF SEQ ID NOS: 5642
/ SOFTWARE: SeqMan99, version 1.03
/ SEQ ID NO 5118
/ LENGTH: 1263
/ TYPE: PRF
/ ORGANISM: Staphylococcus aureus
/ FEATURE:
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (1)..(1263)
/ OTHER INFORMATION: autoLysin [Staphylococcus
```

US-10-471-571A-5118

```
Query Match      3.8%; Score 124; DB 6; Length 1263;
Beet Local Similarity 21.4%; Pred. No. 4.8;
Matches 136; Conservative 64; Mismatches 208; Indels 228; Gaps 35;

QY 79 IYVYNFEYQORAAASVDAKQAVOSLTPGKPYAOGTTDANGVTVQLPKQNGKDAVY 138
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 406 LYDLINEXKLKMGKVAAPWGQSTTPTTPSPKPE----- 438
QY 139 TIKKEPKEG--VVAATNNVAFPVYEMIKQDGSYKYGTEELAVVHIIPKRVNVAND---G 193
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 439 TTPSPKPSYGLKLVANNVA-----QIKPTN-SGLYTT-----VYDKGKATNEVQK 484
QY 194 SLHVKVGTAEENGL-----NGAEF-----VISKEGSPGTVKYIOGVQDS--LYT 237
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 485 TPASKITATLQKQFYLVQDYNQKFKFQVKEGVDVYNTAASPVVNVSYSIKPQTKLYT 544
QY 238 --WTTDKQAKRFITGKSYEIGENDFTEAENGTELTVK-----NLEVGSYILLEVKAP- 289
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 545 VPWGTSKQVA-----GSVSGSNQTFKASKQOQIDKSIYLYGSVNGKS 587
QY 290 ---NNALIEHQTKPTFTIEANNQTPVKEKTVKNDTSKYDKTTPSLDGDVAIGEIKIKYQI 346
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 588 GWVSKAYLVADAKPTP-----TPTPKP-STPTNNKLTVSSLSNG--VA-----QI 629
QY 347 SV---NIPLGIADKEGDANKYV-----KENLYVDKHAALFTDNTVSGEYA 388
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 630 NAKNNGLTFTTYDYDKGKRTKEVQKFAVTAKASLQKGFYLVKQVNSFTLIGWYKQGOVI 689
QY 389 YALYDGDVIAIPENYQVTEQANGFTVAVNPAIYPTLTREGTLKFTYFMHLEKADPTGKF 448
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 690 Y-----NNAKSPVNVQTYTVKKGTLKYSV-----PMGTLY 719
QY 449 KNEANVDNGHTDDQTPP-----TVEVNG--GKRIRIKYD---DVTATQ 487
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 720 KQBAQAVSG-IGNQTFKATKQOQIDKSIYLYGTVNGKSGWSKAYLVAPAKKAVAQPK 778
QY 488 ALAGASFPVRDONSPTJANYLKIDETTKATWKTAAEATTFTTTADGLVDITGLKYG--T 545
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 779 TAVKAYITTKQTTQTVS--KIAQVKPNNYTGIR--ASVYEKTKK-----NGATYADRT 827
QY 546 YTL--EETVAPDDYVLLTNRIEFV-----VNEQSYG---TENLVSPDK----- 584
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 828 FYVTERAHAGNETYVLLANNSTHNIPLGWFNVADLVNQGLKEVKTQKYTVKSNNGLSM 887
QY 585 VP-----NKHGTLPSYG---GKGIYVY 604
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 888 VPWGTKNQVILLTGNNIAGSTFNATKQVSVGKDVLY 923
```

RESULT 15
US-10-471-571A-3542

```
/ Sequence 3542, Application US/10471571A
/ Publication No. US20060115490A1
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SPA
/ TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE: P026927WO
/ CURRENT APPLICATION NUMBER: US/10/471,571A
/ CURRENT FILING DATE: 2003-09-12
/ PRIOR APPLICATION NUMBER: GB-0107661.1
/ PRIOR FILING DATE: 2001-03-27
/ NUMBER OF SEQ ID NOS: 5642
/ SOFTWARE: SeqMan99, version 1.03
/ SEQ ID NO 3542
/ LENGTH: 1629
/ TYPE: PRF
/ ORGANISM: Staphylococcus aureus
/ FEATURE:
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (1)..(1629)
/ OTHER INFORMATION: hypothetical protein
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 02:36:51 ; Search time 50 Seconds
(without alignments)
1097.634 Million cell updates/sec

Title: US-10-661-809a-13

Perfect score: 3225
Sequence: 1 MKQLKKWVTSTLLILPL.....GAVLLIAGVPRARRKENA 627

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:
1: /EMC Celerra_SIDS3/ptodata/2/1aa/5-COMB.pep:*
2: /EMC Celerra_SIDS3/ptodata/2/1aa/6-COMB.pep:*
3: /EMC Celerra_SIDS3/ptodata/2/1aa/7-COMB.pep:*
4: /EMC Celerra_SIDS3/ptodata/2/1aa/H-COMB.pep:*
5: /EMC Celerra_SIDS3/ptodata/2/1aa/PCTUS-COMB.pep:*
6: /EMC Celerra_SIDS3/ptodata/2/1aa/RE-COMB.pep:*
7: /EMC Celerra_SIDS3/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3225	100.0	627	US-09-071-035-218	Sequence 218, App
2	3225	100.0	627	US-10-206-576-218	Sequence 218, App
3	3225	100.0	659	US-09-134-000C-6124	Sequence 6124, App
4	2882	89.7	560	US-09-071-035-220	Sequence 220, App
5	2882	89.7	560	US-10-206-576-220	Sequence 220, App
6	2446	75.8	664	US-09-107-532A-7252	Sequence 7252, App
7	438.5	12.9	477	US-09-134-000C-6123	Sequence 6123, App
8	414.5	12.9	660	US-09-107-532A-6715	Sequence 6715, App
9	395	12.2	508	US-09-107-532A-5331	Sequence 5331, App
10	347	10.8	430	US-09-071-035-274	Sequence 274, App
11	347	10.8	430	US-10-206-576-274	Sequence 274, App
12	311	9.6	384	US-09-071-035-276	Sequence 276, App
13	311	9.6	384	US-10-206-576-276	Sequence 276, App
14	296.5	9.2	665	US-09-769-787-127	Sequence 127, App
15	290.5	9.0	341	US-09-107-532A-5384	Sequence 5384, App
16	232	7.2	2032	US-09-071-035-458	Sequence 458, App
17	232	7.2	2032	US-09-071-035-462	Sequence 462, App
18	232	7.2	2032	US-09-071-035-466	Sequence 466, App
19	232	7.2	2032	US-10-206-576-458	Sequence 458, App
20	232	7.2	2032	US-10-206-576-462	Sequence 462, App
21	232	7.2	2032	US-09-134-000C-6112	Sequence 6112, App
22	219.5	6.8	413	US-09-107-532A-6433	Sequence 6433, App
23	217.5	6.7	621	US-09-071-035-468	Sequence 468, App
24	217.5	6.7	621	US-10-206-576-468	Sequence 468, App
25	201	6.2	1154	US-09-134-000C-6122	Sequence 6122, App

27	197.5	6.1	1161	2	US-09-327-536-2	Sequence 2, App1
28	196.5	6.1	1095	2	US-09-107-532A-3855	Sequence 3855, App
29	193.5	6.0	905	2	US-09-134-000C-4420	Sequence 4420, App
30	193.5	6.0	1112	1	US-08-714-402-2	Sequence 2, App1
31	190.5	5.9	910	2	US-09-134-000C-4288	Sequence 4288, App
32	190	5.9	1027	2	US-09-107-532A-6675	Sequence 6675, App
33	190	5.9	1221	2	US-09-107-532A-3959	Sequence 3959, App
34	189	5.9	688	2	US-09-071-035-464	Sequence 464, App
35	189	5.9	688	2	US-10-206-576-464	Sequence 464, App
36	189	5.9	1074	2	US-09-071-035-358	Sequence 358, App
37	189	5.9	1074	2	US-09-071-035-394	Sequence 394, App
38	189	5.9	1074	2	US-10-206-576-358	Sequence 358, App
39	189	5.9	1074	2	US-10-206-576-394	Sequence 394, App
40	189	5.9	1096	2	US-09-134-000C-5764	Sequence 5764, App
41	184	5.7	952	2	US-09-107-532A-4706	Sequence 4706, App
42	183.5	5.7	390	2	US-09-769-787-128	Sequence 128, App
43	183	5.7	1638	2	US-09-071-035-258	Sequence 258, App
44	183	5.7	1638	2	US-09-071-035-262	Sequence 262, App
45	183	5.7	1638	2	US-09-071-035-266	Sequence 266, App

ALIGNMENTS

RESULT 1
US-09-071-035-218
Sequence 218, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: G11 H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESSES:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brooks
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 218:
SEQUENCE CHARACTERISTICS:
LENGTH: 627 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-218
Query Match 100.0%; Score 3225; DB 2; Length 627;
Best Local Similarity 100.0%; Pred. No. 1.5e-246; Indels 0; Gaps 0;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MKQLKKWVTSTLLILPLFTSVLGTTAFENGESAOVYHKKMTDLPDPLIONSG 60

Db 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPAEENGSAQLVIHKKMTDLPDPLIUNG 60
QY 61 KEMSEFDKYQGLADYTFISIVNTNEFEQRAAGASVDAKQAVOSLTTPKPAQGTDTAN 120
Db 61 KEMSEFDKYQGLADYTFISIVNTNEFEQRAAGASVDAKQAVOSLTTPKPAQGTDTAN 120
QY 121 GNVTVQLPKQNGKDAVTTIKKEPKRGVVAATNMVAPFVYEMIKQTDGSKYGTBEALV 180
Db 121 GNVTVQLPKQNGKDAVTTIKKEPKRGVVAATNMVAPFVYEMIKQTDGSKYGTBEALV 180
QY 181 VHIYKQNVVANDGSLHVKVGTAEENGGLNGAEFVSKSESGPGTVKYIQGVNDGXYTWT 240
Db 181 VHIYKQNVVANDGSLHVKVGTAEENGGLNGAEFVSKSESGPGTVKYIQGVNDGXYTWT 240
QY 241 DKEQAKRFITGSKSYEIGENDFTEAENGTELTVKNLEVGSYLLEEKAPNNAELLINQTK 300
Db 241 DKEQAKRFITGSKSYEIGENDFTEAENGTELTVKNLEVGSYLLEEKAPNNAELLINQTK 300
QY 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIGEKIKYQISVNIPLGIADKGD 360
Db 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIGEKIKYQISVNIPLGIADKGD 360
QY 361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGDVIAPENYQVTEQANGFTVAVNPAY 420
Db 361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGDVIAPENYQVTEQANGFTVAVNPAY 420
QY 421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDQTPPTVEVVTGSKGFIKVD 480
Db 421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDQTPPTVEVVTGSKGFIKVD 480
QY 481 GGVATQALAGASFVVRQNSDPTANYLKIDETTKAATVTKAABATTTTADGLVDITG 540
Db 481 GGVATQALAGASFVVRQNSDPTANYLKIDETTKAATVTKAABATTTTADGLVDITG 540
QY 541 LKGYTYLLEETVAPDDYVLLTNRIEFVNEOSYGTENLVSPKYPNKHGTLPTSGGK 600
Db 541 LKGYTYLLEETVAPDDYVLLTNRIEFVNEOSYGTENLVSPKYPNKHGTLPTSGGK 600
QY 601 IYVYLGSGAVLLIAGVYFARRRKENA 627
Db 601 IYVYLGSGAVLLIAGVYFARRRKENA 627

RESULT 2
US-10-206-576-218
; Sequence 218 Application US/10206576
; Patent No. 6913907
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031

; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB369P1D1
; INFORMATION FOR SEQ ID NO: 218:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 627 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 218:
US-10-206-576-218

Query Match 100.0%; Score 3225; DB 2; Length 627;
Best Local Similarity 100.0%; Pred. No. 1.5e-248;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPAEENGSAQLVIHKKMTDLPDPLIUNG 60
Db 1 MKQLKKWYTVSTLLILPLFTSVLGTTTAPAEENGSAQLVIHKKMTDLPDPLIUNG 60
QY 61 KEMSEFDKYQGLADYTFISIVNTNEFEQRAAGASVDAKQAVOSLTTPKPAQGTDTAN 120
Db 61 KEMSEFDKYQGLADYTFISIVNTNEFEQRAAGASVDAKQAVOSLTTPKPAQGTDTAN 120
QY 121 GNVTVQLPKQNGKDAVTTIKKEPKRGVVAATNMVAPFVYEMIKQTDGSKYGTBEALV 180
Db 121 GNVTVQLPKQNGKDAVTTIKKEPKRGVVAATNMVAPFVYEMIKQTDGSKYGTBEALV 180
QY 181 VHIYKQNVVANDGSLHVKVGTAEENGGLNGAEFVSKSESGPGTVKYIQGVNDGXYTWT 240
Db 181 VHIYKQNVVANDGSLHVKVGTAEENGGLNGAEFVSKSESGPGTVKYIQGVNDGXYTWT 240
QY 241 DKEQAKRFITGSKSYEIGENDFTEAENGTELTVKNLEVGSYLLEEKAPNNAELLINQTK 300
Db 241 DKEQAKRFITGSKSYEIGENDFTEAENGTELTVKNLEVGSYLLEEKAPNNAELLINQTK 300
QY 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIGEKIKYQISVNIPLGIADKGD 360
Db 301 TPFTTEANNQTPVEKTVKNDTSKVDKTPSLDGKQVAIGEKIKYQISVNIPLGIADKGD 360
QY 361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGDVIAPENYQVTEQANGFTVAVNPAY 420
Db 361 ANKYKFNLVDRHDAALTFDNTVSGEYAYALYDGDVIAPENYQVTEQANGFTVAVNPAY 420
QY 421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDQTPPTVEVVTGSKGFIKVD 480
Db 421 IPTLPGGTLKRVYFMHINEKADPTKGFKNBANVDNGHTDQTPPTVEVVTGSKGFIKVD 480
QY 481 GGVATQALAGASFVVRQNSDPTANYLKIDETTKAATVTKAABATTTTADGLVDITG 540
Db 481 GGVATQALAGASFVVRQNSDPTANYLKIDETTKAATVTKAABATTTTADGLVDITG 540
QY 541 LKGYTYLLEETVAPDDYVLLTNRIEFVNEOSYGTENLVSPKYPNKHGTLPTSGGK 600
Db 541 LKGYTYLLEETVAPDDYVLLTNRIEFVNEOSYGTENLVSPKYPNKHGTLPTSGGK 600
QY 601 IYVYLGSGAVLLIAGVYFARRRKENA 627
Db 601 IYVYLGSGAVLLIAGVYFARRRKENA 627

RESULT 3
US-09-134-000C-6124
; Sequence 6124, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

;; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 032796-032
;; CURRENT APPLICATION NUMBER: US/09/134,000C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/055,778
;; PRIOR FILING DATE: 1997-08-15
;; NUMBER OF SEQ ID NOS: 6812
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 6124
;; LENGTH: 659
;; TYPE: PRF
;; ORGANISM: Enterococcus faecalis
;; US-09-134-000C-6124

Query Match 100.0%; Score 3225; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 1.7e-248;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQOLKKWYVTSLLILPLFTSVIGTTAPAEENGESQVLVHKQMTDLPPLIQNSG 60
DB 33 MQOLKKWYVTSLLILPLFTSVIGTTAPAEENGESQVLVHKQMTDLPPLIQNSG 92
QY 61 KEMSEFDKXQGLADVTFSIYVNTNEFEQRAAGASVDAKQAVQSLTPGKPVAGTTDAN 120
DB 93 KEMSEFDKXQGLADVTFSIYVNTNEFEQRAAGASVDAKQAVQSLTPGKPVAGTTDAN 152
QY 121 GNVYVQLPKQNGKQAVYTIKEPEKGGVVAATMMVAFVYEMIKQTDOSYKGTBELAV 180
DB 153 GNVYVQLPKQNGKQAVYTIKEPEKGGVVAATMMVAFVYEMIKQTDOSYKGTBELAV 212
QY 181 VHIYPPNVVANDGSLHVKVGTAEENGELNGAEFVLSKESGPGTVKXIQGVKQGLYTWTT 240
DB 213 VHIYPPNVVANDGSLHVKVGTAEENGELNGAEFVLSKESGPGTVKXIQGVKQGLYTWTT 272
QY 241 DKEQAKRFITGKSYEIGENDPTEAENGTELTIVKQLEVGSYLLEEVKAPNNAELIENQTK 300
DB 273 DKEQAKRFITGKSYEIGENDPTEAENGTELTIVKQLEVGSYLLEEVKAPNNAELIENQTK 332
QY 301 TPFTIEANNQTVETKVDTSKVDKTTPTSLDQKVAIEKIKYQISVNIPLGIADKEED 360
DB 333 TPFTIEANNQTVETKVDTSKVDKTTPTSLDQKVAIEKIKYQISVNIPLGIADKEED 392
QY 361 ANKYKFNVLVDGDAALTPDNVTSGEYAYALYDQTVIAPENYQVTEQANGFTVAVNPAY 420
DB 393 ANKYKFNVLVDGDAALTPDNVTSGEYAYALYDQTVIAPENYQVTEQANGFTVAVNPAY 452
QY 421 IPTLPFGGTLKFVYFMHLEKADPTKGFQKNEAVNDGHTDQTPPTVEVVTGSKREIKYD 480
DB 453 IPTLPFGGTLKFVYFMHLEKADPTKGFQKNEAVNDGHTDQTPPTVEVVTGSKREIKYD 512
QY 481 GNVYVQLPKQNGKQAVYTIKEPEKGGVVAATMMVAFVYEMIKQTDOSYKGTBELAV 540
DB 513 GNVYVQLPKQNGKQAVYTIKEPEKGGVVAATMMVAFVYEMIKQTDOSYKGTBELAV 572
QY 541 LKGYVYEEETAPDDVYVLTNRIFVNVNOSYGTTEMLVSPKVPNKGKGLPTSGK 600
DB 573 LKGYVYEEETAPDDVYVLTNRIFVNVNOSYGTTEMLVSPKVPNKGKGLPTSGK 632
QY 601 IYVYLGSGAVLLLIAGVYFARRRKENA 627
DB 633 IYVYLGSGAVLLLIAGVYFARRRKENA 659

RESULT 4
US-09-071-035-220
Sequence 220, Application US/09071035
Patent No. 6448043

;; GENERAL INFORMATION:
;; APPLICANT: Gail H. Choi
;; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
;; NUMBER OF SEQUENCES: 496
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Human Genome Sciences, Inc.

;; STREET: 9410 Key West Avenue
;; CITY: Rockville
;; STATE: Maryland
;; COUNTRY: USA
;; ZIP: 20850
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
;; COMPUTER: HP Vectra 486/33
;; OPERATING SYSTEM: MSDOS version 6.2
;; SOFTWARE: ASCII Text
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/071,035
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: A. Anders Brookes
;; REGISTRATION NUMBER: 36,373
;; REFERENCE/DOCKET NUMBER: PB369P2
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 220:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 560 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-071-035-220

Query Match 89.7%; Score 2892; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 4.8e-222;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 BENGSAQVLVHKQMTDLPPLIQNSGKEMSEFDKXQGLADVTFSIYVNTNEFEQRA 92
DB 1 BENGSAQVLVHKQMTDLPPLIQNSGKEMSEFDKXQGLADVTFSIYVNTNEFEQRA 60
QY 93 GASVDAKQAVQSLTPGKPVAGTTDANGNVTVOLPKQNGKQAVYTIKEPEKGGVVAAT 152
DB 61 GASVDAKQAVQSLTPGKPVAGTTDANGNVTVOLPKQNGKQAVYTIKEPEKGGVVAAT 120
QY 153 NNVAAPVYEMIKQTDOSYKGTBELAVVHIYPPNVVANDGSLHVKVGTAEENGELNGAE 212
DB 121 NNVAAPVYEMIKQTDOSYKGTBELAVVHIYPPNVVANDGSLHVKVGTAEENGELNGAE 180
QY 213 FVLSKESGPGTVKXIQGVKQGLYTWTTDKQAKRFITGKSYEIGENDPTEAENGTELT 272
DB 181 FVLSKESGPGTVKXIQGVKQGLYTWTTDKQAKRFITGKSYEIGENDPTEAENGTELT 240
QY 273 VQNLVVGSYLLEEVKAPNNAELIENQTKPTFTIEANNQTVETKVDTSKVDKTTPTSLD 332
DB 241 VQNLVVGSYLLEEVKAPNNAELIENQTKPTFTIEANNQTVETKVDTSKVDKTTPTSLD 300
QY 333 GNDVAIGEKIKYQISVNIPLGIADKEGDANKYKFNVLVDGDAALTPDNVTSGEYAYALY 392
DB 301 GNDVAIGEKIKYQISVNIPLGIADKEGDANKYKFNVLVDGDAALTPDNVTSGEYAYALY 360
QY 393 DGDVYIAPENYQVTEQANGFTVAVNPAYIPTLPFGGTLKFVYFMHLEKADPTKGFQKNEA 452
DB 361 DGDVYIAPENYQVTEQANGFTVAVNPAYIPTLPFGGTLKFVYFMHLEKADPTKGFQKNEA 420
QY 453 NVNDGHTDQTPPTVEVVTGSKREIKYDQVYATATAGASFVVDQNSDTANYLKIDET 512
DB 421 NVNDGHTDQTPPTVEVVTGSKREIKYDQVYATATAGASFVVDQNSDTANYLKIDET 480
QY 513 TKAATVYTKAEATFTTTADGLVDITGLKGYVYEEETAPDDVYVLTNRIFVNVNOS 572
DB 481 TKAATVYTKAEATFTTTADGLVDITGLKGYVYEEETAPDDVYVLTNRIFVNVNOS 540

QY 573 YGTTENLVSPKVPKHKGT 592
|||||
Db 541 YGTTENLVSPKVPKHKGT 560

RESULT 5
US-10-206-576-220

; Sequence 220, Application US/10206576
; Patent No. 6913907

GENERAL INFORMATION:

APPLICANT: Choi et al.

TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-R

COMPUTER: Dell Latitude

OPERATING SYSTEM: Windows 98

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/206,576

FILING DATE: 29-Jul-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/071,035

FILING DATE: 1998-05-04

APPLICATION NUMBER: US 60/046,655

FILING DATE: 1997-05-16

APPLICATION NUMBER: US 60/044,031

FILING DATE: 1997-05-06

APPLICATION NUMBER: US 60/066,009

FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:

NAME: Hyman, Mark J.

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB369P1D1

INFORMATION FOR SEQ ID NO: 220:

SEQUENCE CHARACTERISTICS:

LENGTH: 560 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 220:

US-10-206-576-220

Query Match 89.7%; Score 2892, DB 2, Length 560;

Best Local Similarity 100.0%; Pred. No. 4.8e-222;

Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 BENGSAOLVTHKKMTDLPDLIONSGKEMSEPDKYQGLAVTTSIVNVTNEFYBORA 92
|||||

Db 1 BENGSAOLVTHKKMTDLPDLIONSGKEMSEPDKYQGLAVTTSIVNVTNEFYBORA 60
|||||

QY 93 GASVDAKAQAVOSLTGKPEAQTDDANGNVTVOLPKKONGDAVYTIKEBPKEGVAAAT 152
|||||

Db 61 GASVDAKAQAVOSLTGKPEAQTDDANGNVTVOLPKKONGDAVYTIKEBPKEGVAAAT 120
|||||

QY 153 NNVVAFPVYEMTKQTDGSKYGTETELAVVHIIPKVVANDGSLHVKVGTAEENGNGAE 212
|||||

Db 121 NNVVAFPVYEMTKQTDGSKYGTETELAVVHIIPKVVANDGSLHVKVGTAEENGNGAE 180
|||||

QY 213 FVTSSESGPGTVKTIQVKGDLVYTTTDDKQAKRFTGSKYIEGENDTEAENGTELT 272
|||||

Db 181 FVTSSESGPGTVKTIQVKGDLVYTTTDDKQAKRFTGSKYIEGENDTEAENGTELT 240
|||||

QY 273 VKNLEVGSYLIEBVAPNNAELIENQTKPTIEANNQTPVEKTVKNDSKYDKTTPSLD 332
|||||

Db 241 VKNLEVGSYLIEBVAPNNAELIENQTKPTIEANNQTPVEKTVKNDSKYDKTTPSLD 300
|||||

QY 333 GKDVATGSKIKYQISVNIPLGIADKEGDANKYVKNVYDKDAALTENNNTSGEAYALY 392
|||||

Db 301 GKDVATGSKIKYQISVNIPLGIADKEGDANKYVKNVYDKDAALTENNNTSGEAYALY 360
|||||

QY 393 DDDTVIAPENTQVEQANGFTVAVNPAYIPTLTGGLTKFYFMHANEKADPTKGFKNEA 452
|||||

Db 361 DDDTVIAPENTQVEQANGFTVAVNPAYIPTLTGGLTKFYFMHANEKADPTKGFKNEA 420
|||||

QY 453 NVDNHGTDDQTPPEVYVVTGGRPIKVDGVTATQALAGASFVVDONSDTANYIKIDET 512
|||||

Db 421 NVDNHGTDDQTPPEVYVVTGGRPIKVDGVTATQALAGASFVVDONSDTANYIKIDET 480
|||||

QY 513 TKAATWVTKAATFTTTADGVDITGKXGTYLLEETVAPDDVLLTNREFVFNBS 572
|||||

Db 481 TKAATWVTKAATFTTTADGVDITGKXGTYLLEETVAPDDVLLTNREFVFNBS 540
|||||

QY 573 YGTTENLVSPKVPKHKGT 592
|||||

Db 541 YGTTENLVSPKVPKHKGT 560
|||||

RESULT 6
US-09-107-532A-7252
; Sequence 7252, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

CORRESPONDENCE ADDRESS:

ADDRESSER: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 7252:

SEQUENCE CHARACTERISTICS:

LENGTH: 664 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc_feature

LOCATION: (B) LOCATION 1...664

SEQUENCE DESCRIPTION: SEQ ID NO: 7252:

US-09-107-532A-7252

Query Match

75.8%; Score 2446; DB 2; Length 664;

Best Local Similarity 74.2%; Pred. No. 2,4e-186; Matches 462; Conservative 64; Mismatches 95; Indels 2; Gaps 1;

```
QY 1 MQLKVMWTVSTLLILPLFTSVLGTTPAFENEGESQVLVHKKKMTDLPPLIIONSG 60
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 40 MKNHKKINMLGVFLILPLITNSFSAKKVFAEET--AAOVILHKKKMTDLPPLIIONSG 97
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 61 KEMSEPDYQGLADVTFSYVNTNEFYEQRAAGASVDAKQAVQSLLPGKPAVAGTTDAN 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 98 KEMSEPDYQGLADISFVSYNVQEFYARDKGASVDAKQAVQSLLPGKPAVAGTTDAD 157
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 121 GNVTVOLPKKONKDAVTTIKKEPKGCVVAATMVAAPVYEMIKOTDGSYKKGTEELAV 180
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 158 GNVTVSLPKKONKDAVTTIKKEPKGCVSAANMVAAPVYEMIKODDSYKKGTEELPT 217
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 181 VHIYPRNVVANGSLHVKKVGTAEENGANGAFVLSKSGSPGTVKYIQGVQGLYTWTT 240
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 218 IHLVPRNVVANGDGLVTKIQTAEENGANGAFVLSKSGSPGTVKYIQGVQGLYTWTT 277
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 241 DKEQAKRFTGKSYEIGENDPTEAENGTELTVKNLEVGSYLIEVKAAPNNAELIENQTK 300
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 278 DGTAKHFTGHSYDIGNDDFAASTIEKQGLIVNHLVEGKNLEBKAPDNEMIEKQTI 337
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 301 TPFTTEANNQTEVEKTVKNDTSKVDKTPPSLDGKDAVIGBKIKYQISVNIPLGIADKEGD 360
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 338 TPFEILANSQTEPEKTIKNDTSKVDKTPPOLNKGDAVIGBKIQEISVNIPLGIADKEGT 397
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 361 AKRYKENVLDKDALTPDNVTSGEYAVAYLDGTVIAPENYQVTEANGFTVANPAY 420
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 398 QNRKTYTFKIDTHDALTPDNDSSTGYAVALYDGNKEIDPVNSVTEQDGFVSDPVNY 457
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 421 IFTLPGGTLKFEVFMHNEKADPTKGFKNKANVNDGHTDDQTPPYEVVVGSKRFIKYD 480
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 458 ISLTPGGLTKFYVYHNEKADPTKGFKNKANVNDGHTDDQTPPYEVVVGSKRFIKYD 517
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 481 GNVATQALAGASFVVRDONSDTANYLKIDETTKAATVWTKAATFTTTADGLVDITG 540
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 518 GNVTSQTLAGASFVVRDSDPTAKYLSIDPTKAVSWSAKESATVFTTNSGLIDVTG 577
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 541 LKYGTYLLEETVAPDDYVLLTNRIEFVNVNQSSTENLVSPKVNHNKGTLPSTGGG 600
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 578 LKYGTYLLEETVAPDKVPLTNVAFITIDGOSTVYTAGOLISPKIKHNKGTLPSTGGG 637
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 601 IYVYLGSAVLLIAGVFPARR 623
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 638 IYVYLGSAVLLIAGVFPARR 660
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 7

US-09-134-000C-6123

Sequence 6123, Application US/09134000C

Patent No. 6617156

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

FILE REFERENCE: 032796-032

CURRENT APPLICATION NUMBER: US/09/134,000C

PRIOR FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/055,778

NUMBER OF SEQ ID NOS: 6812

SOFTWARE: PatentIn version 3.1

SEQ ID NO 6123

LENGTH: 477

TYPE: PR

ORGANISM: Enterococcus faecalis

US-09-134-000C-6123

Query Match 13.6%; Score 438.5; DB 2; Length 477;

Best Local Similarity 25.9%; Pred. No. 2.1e-26; Matches 166; Conservative 70; Mismatches 211; Indels 193; Gaps 22;

```
QY 8 WTVSTLLILPLFTSVLGTTPAFENEGESQVLVHKKKMTDLPPLIIONSGKEMEB 65
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 7 WLSICWMLALFGFSQ-----QALASASQASQVTLHKKLLFPDGOQLPEOD-QVYGEGLT 60
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 66 PKYQGLADVTFSYVNTNEFYEQRAAGASVDAKQ--AVQSLTPGKPAVAGTTDA----- 119
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 LQNTKGLNDVTVQVNDVTPDYQLRSBEGKTQVBAQQLAETGATNRKPIAEDKTYTINGE 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 120 NGNVTVOLPKK---QNGKDAVTTIKKEPKGCVVAATMVAAPVYEMIKOTDGSYKKGTE 176
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 DGVVSFSLASXOSQGRDKAVLVFEABAAPVNEKKSANLVVILPYDPQSGS----- 171
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 177 ELAVNHIYPRNVVANGSLHVKKVGTAEENGANGAFVLSKSGSPGTVKYIQGVQGLY 236
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 172 -LTHIHLVPRKN----- 181
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 237 TWTTDKEQAKRFTGKSYEIGENDPTEAENGTELTVKNLEVGSYLIEVKAAPNNAELLE 296
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 182 -----EENAYDL----- 188
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 297 NOTKTPTEANNQTEVEKTVKNDTSKVDKTPPSLDGKDAVIGBKIKYQISVNIPLGIAD 356
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 189 -----PLEKTVLDDKQGFNQ-----GEHINQVLTQIPANITG 222
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 357 KEGDANKYKENVLDKDALTPDNVTSGEYAVAYLDGTVIAPENYQVTEANGFTVANPAY 416
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 223 -----YQEFRLSDADKDTLTL---LPRESLEVKAAGKTVTT--GYTLTQHGKFTPLD 269
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 417 NPAVITPLTPGGTLKFEVFMHNEKADPTKGFKNKANVNDGHTDDQTPPYEVVVGSKR 475
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 270 SIKDLQNFRA-NQTMVVSQMRLEKTAEPDTAINNESQVLTDKHT--LTKRATVFGGAS 325
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 476 FIKVDGDAVATQALAGASFVVRDONSDTANYLKIDETTKAATVWTKAATFTTTADGL 535
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 326 FVKVDSR-NAKTLTPPAVPIVKNQAGEYLN-----ETANGYKQKQKALAKKTSQAGE 379
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 536 VDTGLKYGTYLLEETVAPDDYVLLTNRIEFVNVNQSSTENLVSPKVNHNKGTLPSTGG 592
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 380 FSVKGLKQGYFLIEETISAPKGYLNTQETIPIFTVGKRSYATNGQRTAPLHVINKVSESG 439
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 593 LPSTG-GKGIYVYLGSAVLLIAGV-----YFPARRKE 625
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 440 LPTKMEERSIMWTIAG---LTIIGNVVIMLFYQKQKE 475
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 8

US-09-107-532A-6715

Sequence 6715, Application US/09107532A

Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

CORRESPONDENCE ADDRESS: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

ADDRESS: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

```

1      APPLICATION NUMBER: 60/085,598
2      FILING DATE: 14 May 1998
3      APPLICATION NUMBER: 60/051571
4      FILING DATE: July 2, 1997
5      ATTORNEY/AGENT INFORMATION:
6      NAME: Ariniello, Pamela Deneke
7      REGISTRATION NUMBER: 40,489
8      REFERENCE/DOCKET NUMBER: GTC-012
9      TELECOMMUNICATION INFORMATION:
10     TELEPHONE: (781)893-5007
11     TELEFAX: (781)893-8277
12     INFORMATION FOR SEQ ID NO: 6715:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH: 660 amino acids
15     TYPE: amino acid
16     TOPOLOGY: linear
17     MOLECULE TYPE: protein
18     HYPOTHETICAL: YES
19     ORIGINAL SOURCE:
20     ORGANISM: Enterococcus faecium
21     FEATURE:
22     NAME/KEY: misc feature
23     LOCATION: (B) LOCATION 1...660
24     SEQUENCE DESCRIPTION: SEQ ID NO: 6715:
25     US-09-107-532A-6715

```

QY 551 TVAPDDVYLTLNRIIEVNVNEOSYGTENTVSEKXPNKHGKGLPSYGGKGIYY----- 604
 DB 568 TKAFENYOKLTETIEISFYKDKDSY-KEENRIT--IKNNQASAVPMGTSNGFQYVYLISCL 643
 QY 605 -LGSQAVLLLLIAGYFARR 622
 DB 644 LLAGAG---LSAVVYFKKK 659

```

1      RESULT 9
2      US-09-107-532A-5331
3      / Sequence 5331, Application US/09107532A
4      / Patent No. 6583275
5      / GENERAL INFORMATION:
6      / APPLICANT: Lynn A Doucette-Stamm and David Bush
7      / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
8      / ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
9      /
10     / NUMBER OF SEQUENCES: 7310
11     / CORRESPONDENCE ADDRESS:
12     / ADDRESSEE: GENOME THERAPEUTICS CORPORATION
13     / STREET: 100 Beaver Street
14     / CITY: Waltham
15     / STATE: Massachusetts
16     / COUNTRY: USA
17     / ZIP: 02354
18     /
19     / COMPUTER READABLE FORM:
20     / MEDIUM TYPE: CD-ROM ISO9660
21     / COMPUTER: PC
22     / OPERATING SYSTEM: <Unknown>
23     / SOFTWARE: ASCII
24     /
25     / CURRENT APPLICATION DATA:
26     / APPLICATION NUMBER: US/09/107,532A
27     / FILING DATE: 30-Jun-1998
28     / PRIOR APPLICATION DATA:
29     / APPLICATION NUMBER: 60/085,598
30     / FILING DATE: 14 May 1998
31     / APPLICATION NUMBER: 60/051571
32     / FILING DATE: July 2, 1997
33     /
34     / ATTORNEY/AGENT INFORMATION:
35     / NAME: Arinello, Pamela Deneke
36     / REGISTRATION NUMBER: 40,489
37     / REFERENCE/DOCKET NUMBER: GTC-012
38     /
39     / TELECOMMUNICATION INFORMATION:
40     / TELEPHONE: (781)893-5007
41     / TELEFAX: (781)893-8277
42     /
43     / INFORMATION FOR SEQ ID NO: 5331:
44     / SEQUENCE CHARACTERISTICS:
45     / LENGTH: 508 amino acids
46     / TYPE: amino acid
47     / TOPOLOGY: linear
48     / MOLECULE TYPE: protein
49     / HYPOTHETICAL: YES
50     / ORIGINAL SOURCE:
51     / ORGANISM: Enterococcus faecium
52     /
53     / FEATURE:
54     / NAME/KEY: misc feature
55     / LOCATION: (B) LOCATION 1...508
56     /
57     / SEQUENCE DESCRIPTION: SEQ ID NO: 5331:
58     / US-09-107-532A-5331

```


[illegible]

APPLICATION NUMBER: US/10/206,576
FILING DATE: 29-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D1
INFORMATION FOR SEQ ID NO: 274:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 274:
US-10-206-576-274

Query Match 10.8%; Score 347; DB 2; Length 430;
Best Local Similarity 24.0%; Pred. No. 3,6e-19; Indels 186; Gaps 19;
Matches 142; Conservative 64; Mismatches 200;
QY 8 WYVSTLLILPLFTSVLGTTTAFPAENGESAOIVYHKKMTD--LPPLIQNSKENSE 65
DB 6 WLSICVLLALFSGFQ-----QALAEASQASVQVTLHLKLPFGDLPEQO-QNTGEBGL 59
QY 66 FDKYGLADVTFSYVNTNEFYEQRAAGASVDAKO--AVQSITPKGPVAGTTA---- 119
DB 60 LQNYRGINDVTVQVYDVTDVDFYQLRSEKTVQEAQROLAETGATNRKPIAEDKTQTINGE 119
QY 120 NGNVTVOLPKK--QNGKDAVYTIKEPKREGVVAATNMVAVPPVYEMIKOTGSKYGE 176
DB 120 DGVSFSLSKDSQQRDKAYLFFVEADAPFVVEKASNLVILPVDPOGQS----- 170
QY 177 ELAVVHIYPRKVVANDGSLHVKKVGTAEENGNGAEFVLSKESGPGTVKYIQGVKGLY 236
DB 171 -LTHILHYPRK----- 180
QY 237 TWTTDKEQAKRFITGSKSYEIGENDFTEANGTGELTVKNLEVGSYILEEVKAPNNAELIE 296
DB 181 -----EENAYDL----- 187
QY 297 NOTKPTTEIANNQTPVEKTVKNDTSKVDKTPPSLDGKDVAGGEKIKYQISVINIFPIGAD 356
DB 188 -----PLEKTVLDKQOGFRQ-----GEHINYQLTQITANILG 221
QY 357 KEGDANKYKFNLDKDAALTFDNTVSGEYAYALYDGPVTIAPENYQVTEQANGFTYAV 416
DB 222 -----YQEFRISDKADTLTL-----LPESIEVKAAGKTVTT--GYTLTTQKHGFTLDF 268
QY 417 NPAYITPLTGGTLKFVFMHLNEKADPTKGFQNEAN--VDNGHTDDQTPPYEVVTVGGR 475
DB 269 SIKDQNFPA-NQTMVTSYQMRLEKTAEPDTAINNSEQVLTDKGT--LTKRAVTRGGS 324
QY 476 FIKVGDVYATQALAGASFVVDQNSDPTANYLKIDETTKATVWKYKAATFTTTADGL 535
DB 325 FVKVDSF-NAKTLTLEAVPIVKNQAGEYIN-----ETANGYRQKEKALAKFTTSNAGE 378
QY 536 VDIITLKYGTYYLEETVAPDDYVLLTNRILEPVNQSQSYGTENLVSPKRPV 587
DB 379 FSVKXKEMPVLLGRN-----LCTKRLSSESNNRNSFYGKKFLCNERETT 423

RESULT 12
US-09-071-035-276

Sequence 276, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELEPHONE: (301) 309-8512
TELEFAX: (301) 309-8504
INFORMATION FOR SEQ ID NO: 276:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-276

Query Match 9.6%; Score 311; DB 2; Length 384;
Best Local Similarity 23.5%; Pred. No. 2.2e-16;
Matches 127; Conservative 56; Mismatches 180; Indels 178; Gaps 16;
QY 57 QNSGKENSEPDKYGLADVTFSYVNTNEFYEQRAAGASVDAKO--AVQSITPKGPVAV 114
DB 7 QNTGEBGLQNYRGINDVTVQVYDVTDVDFYQLRSEKTVQEAQROLAETGATNRKPIAE 66
QY 115 GTTDA---NGNVTVOLPKK--QNGKDAVYTIKEPKREGVVAATNMVAVPPVYEMIKOT 167
DB 67 DKTQRTINEDGVVSFSLSKDSQQRDKAYLFFVEADAPFVVEKASNLVILPVDPOGQS 126
QY 168 DGSYKGTIEELAVVHIYPRKVVANDGSLHVKKVGTAEENGNGAEFVLSKESGPGTVKY 227
DB 127 -----LTHILHYPRK----- 136
QY 228 IQGVKGLYTTWTDKEQAKRFITGSKSYEIGENDFTEANGTGELTVKNLEVGSYILEEVK 287
DB 137 -----EENAYDL----- 143
QY 288 APNNAELIENQTKPTTEIANNQTPVEKTVKNDTSKVDKTPPSLDGKDVAGGEKIKYQIS 347
DB 144 -----PLEKTVLDKQOGFRQ-----GEHINYQLT 168
QY 348 VNIPIGLDKGDKANKYKFNLDKDAALTFDNTVSGEYAYALYDGPVTIAPENYQVTE 407
DB 169 TQIPANILG-----YQEFRISDKADTLTL-----LPESIEVKAAGKTVTT--GYTLTT 215
QY 408 QANGFTVAVNPAYITPLTPGGTLKFVFMHLNEKADPTKGFQNEAN--VDNGHTDDQTPPT 466
DB 216 QKHGFTLPSIKDQNFPA-NQTMVTSYQMRLEKTAEPDTAINNSEQVLTDKGT--LTKR 271

Qy 467 VEVVVGKRFIKVDGVTATQALAGASFWVRDQNSDTANYLKIDETTKAATVWTKAEBAT 526
| : : : : :
Db 272 ATVRGSGSFVAVDSE-NAKITLPEAVFIVKQAGEYLN-----ETANGYRQKKEKALAK 325
| : : : : :
Qy 527 TTTTADGLVDITGLKGYTYLAEFTVAPDDVLLTNRIEFVNVESYGTTENLVSPKVP 586
| : : : : :
Db 326 KETSNQAGEFVSVKGKKRWPVLLGRN-----LCTKRLLSSSENRNSFYGKKFKLCNERTT 378
| : : : : :
Qy 587 N 587
| : : : : :
Db 379 N 379

RESULT 13
US-10-206-576-276
; Sequence 276, Application US/10206576
; Patent No. 6913907
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
; FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB369P1D1
; INFORMATION FOR SEQ ID NO: 276:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 276:
US-10-206-576-276

Query Match 9.6%; Score 311; DB 2; Length 384;
Best Local Similarity 23.5%; Pred. No. 2,2e-16;
Matches 127; Conservative 56; Mismatches 180; Indels 178; Gaps 16;

Db 127 -----LTHILYPRN----- 136
| : : : : :
Qy 228 IGVVDGLYTTTDBEQAKRFITGKSYRIGENDPTEANGTGLVYKNLEVGSYILLEBK 287
| : : : : :
Db 137 -----EENAYDL----- 143
| : : : : :
Qy 288 APNNALLEINQKTPFTIEANNQTPVEKTVKNDTSKVDKTPTSLODKOVAIGEKIKYQIS 347
| : : : : :
Db 144 -----PPEKTVLDKQGFNO-----GEHINYQAT 168
| : : : : :
Qy 348 VNIPLGIADKSGDANKYVKNLVDKDAALTPDNVTSGEYAYALVDGDTVIAPENYQTE 407
| : : : : :
Db 169 TQIPANILG-----YQEFRISDKADPTLLT-----LPSIEIVKVAQKVTVT--GYTLTT 215
| : : : : :
Qy 408 QANGFTVAVNPAYITTLTPGGTLKAVYFMHLNEKADPTKGFQKQEAN--VNGHTTDDQTPPT 466
| : : : : :
Db 216 QHGGFTLDFPSIDQLNFA--NQTMTVSYORLEKTAEPDTAINNQSOLVTDKHT--LTKR 271
| : : : : :
Qy 467 VEVVVGKRFIKVDGVTATQALAGASFWVRDQNSDTANYLKIDETTKAATVWTKAEBAT 526
| : : : : :
Db 272 ATVRGSGSFVAVDSE-NAKITLPEAVFIVKQAGEYLN-----ETANGYRQKKEKALAK 325
| : : : : :
Qy 527 TTTTADGLVDITGLKGYTYLAEFTVAPDDVLLTNRIEFVNVESYGTTENLVSPKVP 586
| : : : : :
Db 326 KETSNQAGEFVSVKGKKRWPVLLGRN-----LCTKRLLSSSENRNSFYGKKFKLCNERTT 378
| : : : : :
Qy 587 N 587
| : : : : :
Db 379 N 379

RESULT 14
US-09-769-787-127
; Sequence 127, Application US/09769787
; Patent No. 6936252
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hambro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21128WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 127
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-127

Query Match 9.2%; Score 296.5; DB 2; Length 665;
Best Local Similarity 23.0%; Pred. No. 7.6e-15;
Matches 176; Conservative 84; Mismatches 259; Indels 247; Gaps 38;

Db 146 SLSTVVGSDGALITGSKAVPIEILPLNDVVAHYPPKTEAKPKIDKPKKANPDP1P 205
QY 197 VKKGTAEENEGANGAEF-VISKSBSPGTVKTIQGVKGLYWTDTKQAKFPITGSKYE 255
Db 206 VDKDTPVNHQVGDVVEIYVTK--IPALANATA-----SDMTGLAF 249
QY 256 IGENDFTEANGTGLTYKN--LEVGSYLIEVK-----APNNAELINQOTY 301
Db 250 -----NKGTAVYVDVALLEAGDVALTVAAGFPLKLTDAGLAKKNDQNAEKTYKI 300
QY 302 PFTTEANNQTPVEKTVKD-----TSKVDKTPSLDKDVAIGEKIKYQISV 348
Db 301 TYSATLNDKXALIVEESNDVTFNNGNPDHGNTPKPNK--PNENG-DLTLTKTWDAIGA 357
QY 349 NIPGLIADKEGDANKYKFNVDKIDALTEP--NTSGEYVYALYDGDVTVAEPNYQVT 406
Db 358 PIPAG-----AEATFDLVNAQTKGVQTV---TLTTDRN---T 389
QY 407 EQANGFTVAVNPAYIPTLPGLTKFVYFMHLEKADPTKGFKNENAVNDGHTDDOTPT 466
Db 390 VTVNGLDKTEKTFPERSIKGISADYQELTTAGEIA--VKWKDE---NKPFLDPTBP- 442
QY 467 VEVVTGKRFRIV--DQVYATQALAGASFVVDQNSDTA--NYL--KIDE-----T 512
Db 443 -KVVTYGGKFKVKNKD-----NRLAGAEFVI--ANADNAGQYLARRADKVSQEEKOLVVT 495
QY 513 TK-----AATV--KTAKAATTF 528
Db 496 TKDALDRAVAAVNALTAQOQTOEKEKVDKAAVNAVIAANNAFEWADNDONENVKL 555
QY 529 TTTADGLVDITGLKGTYYLEETVAPDDVLLTNRLEFVNQSYGTENLV----- 580
Db 556 VSDAGREBITLGLTYLEETKOPAGALITSKQKEVTAITSATGQGIIEYTAGSK 615
QY 581 -SPEKVPNGHKGTLPSGTGKIYVYLGSAVLLLIAGVYFARRKE 625
Db 616 DDATKVNK-KITIPOTGGITIIFAVAGAIIMGIAVYVYVKNND 660

RESULT 15
US-09-107-532A-5384
; Sequence 5384, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Atinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 893-5007
; FAX: (781) 893-8277
; INFORMATION FOR SEQ ID NO: 5384:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 341 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEITICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1..341
; SEQUENCE DESCRIPTION: SEQ ID NO: 5384:
US-09-107-532A-5384
Query Match 9.0%; Score 290.5; DB 2; Length 341;
Best Local Similarity 27.1%; Pred. No. 8e-15;
Matches 90; Conservative 58; Mismatches 125; Indels 59; Gaps 13;
335 DVAIGEKIKYQISVNIPLGIADKSGDANKYK-FVLDKIDALTFDNTSGEYVYALYD 393
Db 23 DLERKKTASYTTIAPIDPYFIDSVLENGSAVINKYKITDTPVGLTY-----YD 70
QY 394 -----GDTVIAP-ENYQVTEOANGFTVAV-----NPAYIPTL-----TPGTLKFYF 435
Db 71 QEIVRAGETILTQODIIVEVNSGVYVITLITEENGVAKYD/LGR/LADAGGDLITRYN 130
QY 436 MHLNEKADPTKGFKNENAVNDGHTD-----DQTPTEVEVTGKRFRIVDGDVATQALA 490
Db 131 LKVTLEADPFHNTAVIEIGRNDPEYERKEVPEKVTGGKRFERY--DASSSELK 188
QY 491 GASFVVRQNSDTANY-----LKIDETTKATWKTAEATFTTTADGLVDITGL 541
Db 189 DARFEL--WNDRSEYALFYGESPLAYESGADRIEATSGQATEFVADGNGYFVQGL 246
QY 542 KGTYYLEETVAPDDVLLTNR---IEFVNQSYGTENLV-----SPEKVPNGHKGTL 593
Db 247 DYGTYQMETNAPEBIVLPTGEAATFEIISGISNEIRIQIVGNPPEVYVNNKRGSL 306
QY 594 PSTGKGIYVYLGSAVLLLIAGVYFARRKE 625
Db 307 PATGNGHLAFLILGISIMI--GAYSRYRSK 336

Search completed: July 6, 2006, 02:38:16
Job time : 52 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 18:08:58 ; Search time 7250 Seconds

(without alignments) 16591.048 Million cell updates/sec

Title: US-10-661-809a-12

Sequence: 1 atgaagcaatcaaaaaaagt.....gacgtagaagaagaatgct 1881

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_sy:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_hcg:*
13: gb_in:*
14: gb_cm:*
15: gb_da:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1881	100.0	1980	2	AR396704 Sequence
2	1881	100.0	1983	2	BD191807 Enterococ
3	1881	100.0	1983	2	AR228552 Sequence
4	1881	100.0	1983	2	AR692967 Sequence
5	1881	100.0	15614	2	BD193451 Enterococ
6	1859	98.8	1875	2	CQ969453 Sequence
7	1859	98.8	110000	15	AE016830_10 Continuation (11 o
8	1687	89.7	1687	2	BD191808 Enterococ
9	1687	89.7	1687	2	AR228553 Sequence
10	1687	89.7	1687	2	AR692968 Sequence
11	1014.6	53.9	1878	2	CQ969772 Sequence
12	1014.6	53.9	1995	2	AR448987 Sequence
13	595	31.6	2199	2	AX476900 Sequence
14	94.6	5.0	150038	12	AC181263 Strongylo
15	94.6	5.0	204539	12	AC178783 Strongylo
16	83.8	4.5	2069	2	E10125 DNA encodin
17	83.8	4.5	3399	2	E10126 DNA encodin
18	83.4	4.4	223469	12	AC112451 Rattus no

19	80	4.3	1701	2	CS075468 Sequence
20	80	4.3	1701	2	CS082924 Sequence
21	80	4.3	253001	13	AE014834 Plasmodiu
22	79.2	4.2	198224	6	AC149597 Mus muscu
23	77.4	4.1	153266	6	AC165336 Mus muscu
24	76.2	4.1	129295	12	AC178570 Strongylo
25	74.6	4.0	7218	2	I66494 Sequence 14
26	74.6	4.0	16377	13	AF0202180 Plasmodiu
27	74.4	4.0	233713	12	AC133034 Rattus no
28	73.6	3.9	111554	4	AP002460 Arabidops
29	73	3.9	271546	13	AE014843 Plasmodiu
30	72.2	3.8	5420	13	AF056936 Plasmodiu
31	71	3.8	110000	15	CP000061_3 Continuation (4 of
32	71	3.8	110000	15	CP000061_4 Continuation (5 of
33	71	3.8	175544	12	AC117342 Rattus no
34	70.6	3.8	66442	13	PFMAL1P4 Arabidops
35	70.4	3.7	1686	2	E08995 DNA encodin
36	70	3.7	110000	15	CP000111_04 Continuation (5 of
37	70	3.7	257757	13	AE014837 Plasmodiu
38	69.8	3.7	110000	13	PFMAL1P2_0 Arabidops
39	69.8	3.7	348174	13	CR382399 Plasmodiu
40	69.6	3.7	208332	6	AC122227 Mus muscu
41	69.4	3.7	206187	12	AC182450 Mus muscu
42	69.2	3.7	825	13	AF206632 Plasmodiu
43	69.2	3.7	81019	4	AB073159 Arabidops
44	69.2	3.7	253924	13	AE014822 Plasmodiu
45	68.6	3.6	4995	13	PFMMSA M69183 Plasmodiu

ALIGNMENTS

RESULT 1
AR396704 1980 bp DNA linear PAT 18-DEC-2003
Sequence 2719 from patent US 6617156.
DEFINITION AR396704
ACCESSION AR396704
VERSION AR396704.1 GI:40126444
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1980)
Doucette-Stamm, L.A. and Bush, D.
Nucleic acid and amino acid sequences relating to Enterococcus
faecalis for diagnostics and therapeutics
JOURNAL
Patent: US 6617156-A 2719 09-SBP-2003;
FEATURES
location/Qualifiers
1..1980
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1881; DB 2; Length 1980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGCAATTAATAAAAGTTGGTACACCGTTAGCTTGTACTAATTTGCCACTT 60
97 ATGAAGCAATTAATAAAAGTTGGTACACCGTTAGCTTGTACTAATTTGCCACTT 156
QY 61 TTCACAGCTGTTTGGGACACCACTGATTTGGAGAAAGGAGGAGCGACAG 120
157 TTCACAGCTGTTTGGGACACCACTGATTTGGAGAAAGGAGGAGCGACAG 216
QY 121 CTCGTGATTCACAAAGAAATGACGATTTACCAAGATCCGCTTATCAAAATAGCGGG 180
217 CTCGTGATTCACAAAGAAATGACGATTTACCAAGATCCGCTTATCAAAATAGCGGG 276
QY 181 AAAAGAAATGACGATTTGATTAATATCAAGACATGAGATGATGACGTTTGTATTTAT 240
277 AAAAGAAATGACGATTTGATTAATATCAAGACATGAGATGATGACGTTTGTATTTAT 336
DB

241 AACGTGACGAAGATTTTACGAGCAACGAGCGGACGAGCGCAAGCGTTGATGCACTAAA 300
|
|
|
Db 337 AACGTGACGAAGATTTTACGAGCAACGAGCGGACGAGCGCAAGCGTTGATGCACTAAA 396
|
|
|
Qy 301 CAAGCTGTCGAAGGTTTAACTCCGCGGAAAACCTGTTGCTCAAGAAACCAACGATGCAAA 360
|
|
|
Db 397 CAAGCTGTCGAAGGTTTAACTCCGCGGAAAACCTGTTGCTCAAGAAACCAACGATGCAAA 456
|
|
|
Qy 361 GGAATATGCACTGTTACGTTAAGTAAAGAAATGTAAGATGCAAGTATACCAAT 420
|
|
|
Db 457 GGAATATGCACTGTTACGTTAAGTAAAGAAATGTAAGATGCAAGTATACCAAT 516
|
|
|
Qy 421 AAAGAAAGAACCAAAAGAGGTTGATGCTGCTACCAATATGAGTGGCGTTCCCAAT 480
|
|
|
Db 517 AAAGAAAGAACCAAAAGAGGTTGATGCTGCTACCAATATGAGTGGCGTTCCCAAT 576
|
|
|
Qy 481 TACGAATGATCAAGCAAAACAGATGTTCTTATTAATATGAAACAGAAAGATTAAGCGTT 540
|
|
|
Db 577 TACGAATGATCAAGCAAAACAGATGTTCTTATTAATATGAAACAGAAAGATTAAGCGTT 636
|
|
|
Qy 541 GTTCAATTTATCTTAAAGATGTTAGGCAATGATGTTAGTTACATGTAAGAAAGTA 600
|
|
|
Db 637 GTTCAATTTATCTTAAAGATGTTAGGCAATGATGTTAGTTACATGTAAGAAAGTA 696
|
|
|
Qy 601 GGAATGCTGAAGATGAGATTTAAATGCGCAGAAATTTGTTATTTCTAAAGCGAAGGC 660
|
|
|
Db 697 GGAATGCTGAAGATGAGATTTAAATGCGCAGAAATTTGTTATTTCTAAAGCGAAGGC 756
|
|
|
Qy 661 TCAACGAGCAAGTAAATATATATCCAGAGATCAAGATGATTAATATCATGAGCAACG 720
|
|
|
Db 757 TCAACGAGCAAGTAAATATATATCCAGAGATCAAGATGATTAATATCATGAGCAACG 816
|
|
|
Qy 721 GATTAAGAACCAAGCAAAAGCGTTTATTAATGGAAGAAATTAATGGAATTTGCGAAATGAT 780
|
|
|
Db 817 GATTAAGAACCAAGCAAAAGCGTTTATTAATGGAAGAAATTAATGGAATTTGCGAAATGAT 876
|
|
|
Qy 781 TTCAAGAGAGCAGAGATGAAACGCGGAGAAATTAACAGTTAAATCTTGAGGTGGTTCG 840
|
|
|
Db 877 TTCAAGAGAGCAGAGATGAAACGCGGAGAAATTAACAGTTAAATCTTGAGGTGGTTCG 936
|
|
|
Qy 841 TATATTTTGAAGAGATTAAGTCCAAATTAATGCAATTAATGAAATTAATCAACAAA 900
|
|
|
Db 937 TATATTTTGAAGAGATTAAGTCCAAATTAATGCAATTAATGAAATTAATCAACAAA 996
|
|
|
Qy 901 ACACCATTTTACATTTGAAGCAAAACATCAAAACCTGTTGAAAAAAACAGTCAAAAATGAT 960
|
|
|
Db 997 ACACCATTTTACATTTGAAGCAAAACATCAAAACCTGTTGAAAAAAACAGTCAAAAATGAT 1056
|
|
|
Qy 961 ACCTCTAAAGTTGATTAACCAACACCAAGCTTAGATGTTAAAGATGTGCGCAATTTGCGCA 1020
|
|
|
Db 1057 ACCTCTAAAGTTGATTAACCAACACCAAGCTTAGATGTTAAAGATGTGCGCAATTTGCGCA 1116
|
|
|
Qy 1021 AAAATTAATTAATCAAAATTTCTGTAATATTTCCATTGGGAGATTTGCGAACAAGAGCGAC 1080
|
|
|
Db 1117 AAAATTAATTAATCAAAATTTCTGTAATATTTCCATTGGGAGATTTGCGAACAAGAGCGAC 1176
|
|
|
Qy 1081 GCTAATTAATTAATCGTCAATTTCAATTTAGTTGAATTAACATGACAGCTTTAATCTTTGAT 1140
|
|
|
Db 1177 GCTAATTAATTAATCGTCAATTTCAATTTAGTTGAATTAACATGACAGCTTTAATCTTTGAT 1236
|
|
|
Qy 1141 AACGTGACTTTCGAGAGATATGCTTAATGCGTTATATGATGGGATATCAGTGAATGCTCT 1200
|
|
|
Db 1237 AACGTGACTTTCGAGAGATATGCTTAATGCGTTATATGATGGGATATCAGTGAATGCTCT 1296
|
|
|
Qy 1201 GAAATTTATCAAGTATGATGTAACAAGCAATGCTTCACTGCGCCTTAATCAAGGTAT 1260
|
|
|
Db 1297 GAAATTTATCAAGTATGATGTAACAAGCAATGCTTCACTGCGCCTTAATCAAGGTAT 1356
|
|
|
Qy 1261 ATTCTTACGCTTAAACCAAGCGGCAACATTAATTTGTTTACTTTATGCAATTTAAATGAA 1320
|
|
|
Db 1357 ATTCTTACGCTTAAACCAAGCGGCAACATTAATTTGTTTACTTTATGCAATTTAAATGAA 1416
|
|
|
Qy 1321 AAAGCAGATCTTACGAAGAGCTTTAAATAATGAGGCGAATGTTGATTAACGCTCAATCCGAC 1380

1417 AAAGCAGATCTTACGAAGAGCTTTAAATAATGAGGGAATGTTGATTAAGCTCAATCCGAC 1476
|
|
|
Db 1381 GACCAAAACCAACCAACTGTTGAAGTTTGAACAGTGGGAAAACGTTTCAATTAAGTCAT 1440
|
|
|
Qy 1477 GACCAAAACCAACCAACTGTTGAAGTTTGAACAGTGGGAAAACGTTTCAATTAAGTCAT 1536
|
|
|
Db 1441 GGCAGATGTAAGCAGCAACAACCTTTGGGAGGCTTCTTTGTCGTCGATCAAAAC 1500
|
|
|
Qy 1537 GGCAGATGTAAGCAGCAACAACCTTTGGGAGGCTTCTTTGTCGTCGATCAAAAC 1596
|
|
|
Db 1501 AGCGACACGCAAAATTTATGAAATTCATGAAACACGAAGCAGCAACTTGGGTGAAA 1560
|
|
|
Qy 1597 AGCGACACGCAAAATTTATGAAATTCATGAAACACGAAGCAGCAACTTGGGTGAAA 1656
|
|
|
Db 1561 ACMAAGCTGAAGCAACTTCTTTCAACAAGGCTGATGATTAATGATATCAACAGG 1620
|
|
|
Qy 1657 ACMAAGCTGAAGCAACTTCTTTCAACAAGGCTGATGATTAATGATATCAACAGG 1716
|
|
|
Db 1621 CTTAAATACGTTACCTAATTTATGAAGAACTGTAGCTCTGATGATTAATGCTTTGTA 1680
|
|
|
Qy 1717 CTTAAATACGTTACCTAATTTATGAAGAACTGTAGCTCTGATGATTAATGCTTTGTA 1776
|
|
|
Db 1681 ACAAAATCGAATTTGATTTGTCATGAAACAATCATATGSCAACACAGAAAACCTAGTT 1740
|
|
|
Qy 1777 ACAAAATCGAATTTGATTTGTCATGAAACAATCATATGSCAACACAGAAAACCTAGTT 1836
|
|
|
Db 1741 TCAACGAAAAAGTACCAACCAACAAAGGTACTTAACCTTCAACAGTGGCAAGGA 1800
|
|
|
Qy 1837 TCAACGAAAAAGTACCAACCAACAAAGGTACTTAACCTTCAACAGTGGCAAGGA 1896
|
|
|
Db 1801 ATCTACGTTTACTTGAAGTGGCGCAGTCTTGTACTTAATTCAGAGTCTTACTTGTCT 1860
|
|
|
Qy 1897 ATCTACGTTTACTTGAAGTGGCGCAGTCTTGTACTTAATTCAGAGTCTTACTTGTCT 1956
|
|
|
Db 1861 AGACGTAGAAAAAGAAATGCT 1881
|
|
|
Qy 1957 AGACGTAGAAAAAGAAATGCT 1977
|
|
|

RESULT 2
BD191807 1983 bp DNA linear PART 17-JUL-2003
LOCUS BD191807
DEFINITION Enterococcus Faecalis polynucleotides and polypeptides.
ACCESSION BD191807.1 GI:33001546
VERSION BD191807.1
KEYWORDS JP 2002516571-A/109.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1983)
AUTHORS Kunesh, C.A., Choi, G.H., Bailey, C. and Hromockyj, A.
TITLES Enterococcus Faecalis polynucleotides and polypeptides
JOURNAL Patent: JP 2002516571-A 109 04-JUN-2002;
HUMAN GENOME SCIENCES INC
COMMENT
PN JP 2002516571-A/109
PD 04-JUN-2002
PF 04-MAY-1998 JP 1998548289
PR 06-MAY-1997 US 60/044031, 16-MAY-1997 US 60/046655 PR
14-NOV-1997 US 60/066009
PI CHARLES A KUNESH, GIL H CHOI, CAMELLA BAILEY, ALEX HROMOCKYJ PC
C12N1/5.31, C07K14/315, C07K16/12, C12Q1/68, C12N1/21, C12N5/12, PC
G01N33/569.
PC G01N33/68, A61K39/09
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source Location/Qualifiers
1..1983
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN

Query Match 100.0%; Score 1881; DB 2; Length 1983;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGCAATTAATAAAGTTTGTATACACCGTTAGTACCTGTATTAATTTTGGCACTT 60
 100 ATGAGCAATTAATAAAGTTTGTATACACCGTTAGTACCTGTATTAATTTTGGCACTT 159
 61 TTCAACAAGTGTATTAGGCAACAACCTGCAATTTGCAGAGAAATGGGAGAGCGCAG 120
 160 TTCAACAAGTGTATTAGGCAACAACCTGCAATTTGCAGAGAAATGGGAGAGCGCAG 219
 121 CTCGATTTCACAAAAAGAAATGACGATTTTACCGATCCGCTTATTAATAATGGCGG 180
 220 CTCGATTTCACAAAAAGAAATGACGATTTTACCGATCCGCTTATTAATAATGGCGG 279
 181 AAAAGAAATGACGATTTTATCAAGGACTGGCAGATGTGACGTTTATTTAT 240
 280 AAAAGAAATGACGATTTTATCAAGGACTGGCAGATGTGACGTTTATTTAT 339
 241 AACGTGACGAACTTTTACGAGCAACGAGCGGCAAGCGTGTATGCACTTAA 300
 340 AACGTGACGAACTTTTACGAGCAACGAGCGGCAAGCGTGTATGCACTTAA 399
 301 CAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGTGTCTCAAGAAACCAAGTGCAT 360
 400 CAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGTGTCTCAAGAAACCAAGTGCAT 459
 361 GGGAAATGCTCACTGTCACTTAACTTAAACAAATGTTAAAGATGCACTGTATACAT 420
 460 GGGAAATGCTCACTGTCACTTAACTTAAACAAATGTTAAAGATGCACTGTATACAT 519
 421 AAAAGAAACCAAAAGAGGTGTAGTGTCTGACGAATATGATGTGCGTTCCAGTT 480
 520 AAAAGAAACCAAAAGAGGTGTAGTGTCTGACGAATATGATGTGCGTTCCAGTT 579
 481 TAACGAATGATCAAGCAACAGATGTTCTTAAATATGAAACGAAGAAATTAACGCTT 540
 580 TAACGAATGATCAAGCAACAGATGTTCTTAAATATGAAACGAAGAAATTAACGCTT 639
 541 GTTCATATTTATCTTAAATATGTTAGTACCAATGATGTTTACATGTGAAAAAGTA 600
 640 GTTCATATTTATCTTAAATATGTTAGTACCAATGATGTTTACATGTGAAAAAGTA 699
 601 GGAACCTGTGAAATGAGGATTTAAATGCGCAGAAATTTGTTATTTCTTAAAGCGAAGC 660
 700 GGAACCTGTGAAATGAGGATTTAAATGCGCAGAAATTTGTTATTTCTTAAAGCGAAGC 759
 661 TCACCAAGCAAGTAAATATATCCAGAGATCAAGATGATTTATATCATGACACAGC 720
 760 TCACCAAGCAAGTAAATATATCCAGAGATCAAGATGATTTATATCATGACACAGC 819
 721 GATAAAGAAACGAAGAAAGCTTTATTAATGAGGAAAGTTATGAAATTTGGCAAAATGAT 780
 820 GATAAAGAAACGAAGAAAGCTTTATTAATGAGGAAAGTTATGAAATTTGGCAAAATGAT 879
 781 TTCAACAAGGAGAGATGAAACGAGGAGATTTAAGATTTAAATCTTGAGGTTGCTTCG 840
 880 TTCAACAAGGAGAGATGAAACGAGGAGATTTAAGATTTAAATCTTGAGGTTGCTTCG 939
 841 TATATTTTGAAGAGATTAAGCTCAAAATATGAGAAATTAATGAAATTAATCAACAAA 900
 940 TATATTTTGAAGAGATTAAGCTCAAAATATGAGAAATTAATGAAATTAATCAACAAA 999
 901 ACACCAATTTACATTTGAAGCAACATCAACCTGTGTTGAAAAAAGTCAAAATGAT 960
 1000 ACACCAATTTACATTTGAAGCAACATCAACCTGTGTTGAAAAAAGTCAAAATGAT 1059
 961 ACCTTAAAGTTGATTAACCAACCAAGCTTATGATGTTAAAGATGTCATTTGGCGAA 1020
 1060 ACCTTAAAGTTGATTAACCAACCAAGCTTATGATGTTAAAGATGTCATTTGGCGAA 1119

QY 1021 AAAATTAATATCAATTTCTGTAATAATTCATTTGGGATTTGCAGACAAAGAGCGAC 1080
 DB 1120 AAAATTAATATCAATTTCTGTAATAATTCATTTGGGATTTGCAGACAAAGAGCGAC 1179
 QY 1081 GCTAATTAATATCGTCAAAATTTCAATTTAGTTGATTAACAATGATGACGCTTAACTTTGAT 1140
 DB 1180 GCTAATTAATATCGTCAAAATTTCAATTTAGTTGATTAACAATGATGACGCTTAACTTTGAT 1239
 QY 1141 AACGTGACTTCTGAGAGATATGCTTATATGATGGGATATCAATGATTTGCTCCT 1200
 DB 1240 AACGTGACTTCTGAGAGATATGCTTATATGATGGGATATCAATGATTTGCTCCT 1299
 QY 1201 GAAAATTAATCAAGTATGTAACAAAGGCTTCACTGTGCGGTTAATCAAGGTAT 1260
 DB 1300 GAAAATTAATCAAGTATGTAACAAAGGCTTCACTGTGCGGTTAATCAAGGTAT 1359
 QY 1261 ATTCTACGCTTAAACACGAGCGGCACTAAATTCGTTTACTTTATGCACTTTAAATGAA 1320
 DB 1360 ATTCTACGCTTAAACACGAGCGGCACTAAATTCGTTTACTTTATGCACTTTAAATGAA 1419
 QY 1321 AAAGCAATCTTACGAAAGGCTTTAAATATGAGGCAATGTTGATACGTCATACGAC 1380
 DB 1420 AAAGCAATCTTACGAAAGGCTTTAAATATGAGGCAATGTTGATACGTCATACGAC 1479
 QY 1381 GACCAAAACACCAACGTTGATGATGACAGGATGGGAAACGTTTCAATTAAGTGCAT 1440
 DB 1480 GACCAAAACACCAACGTTGATGATGACAGGATGGGAAACGTTTCAATTAAGTGCAT 1539
 QY 1441 GCGATGTCACAGCAACAGCCTTGGCGGAGCTTCTTGTGCTCCGATCAAAAC 1500
 DB 1540 GCGATGTCACAGCAACAGCCTTGGCGGAGCTTCTTGTGCTCCGATCAAAAC 1599
 QY 1501 AGCGACACGCAATTTATTTGAAATTCATGAAACAGAAAGCAGCACTTGGGTGAAA 1560
 DB 1600 AGCGACACGCAATTTATTTGAAATTCATGAAACAGAAAGCAGCACTTGGGTGAAA 1659
 QY 1561 ACAAAAGCTGAAGCAACTTCTTCAACCAAGCGGATGATGTTGATATCAAGG 1620
 DB 1660 ACAAAAGCTGAAGCAACTTCTTCAACCAAGCGGATGATGTTGATATCAAGG 1719
 QY 1621 CTTAATATCGGTACCTTATTTATGAGAAACCTGAGCTCCTGATGATTTATCTTTGTTA 1680
 DB 1720 CTTAATATCGGTACCTTATTTATGAGAAACCTGAGCTCCTGATGATTTATCTTTGTTA 1779
 QY 1681 ACAATTCGATTTGATTTGTGCTCAATGACATATGCGCAACAGAAACCTTATGTT 1740
 DB 1780 ACAATTCGATTTGATTTGTGCTCAATGACATATGCGCAACAGAAACCTTATGTT 1839
 QY 1741 TCACCAAGAAAGTACCAACCAACCAAGGATCTTACCTTCAACAGTGGCAAGGA 1800
 DB 1840 TCACCAAGAAAGTACCAACCAACCAAGGATCTTACCTTCAACAGTGGCAAGGA 1899
 QY 1801 ATCTACGTTTACTTGAAGTGGCGCAGTCTTGTACTTATTTGACAGAGTCTTACTTGT 1860
 DB 1900 ATCTACGTTTACTTGAAGTGGCGCAGTCTTGTACTTATTTGACAGAGTCTTACTTGT 1959
 QY 1861 AGACGTAGAAAAGAAATGCT 1881
 DB 1960 AGACGTAGAAAAGAAATGCT 1980

RESULT 3
 AR228552 1983 bp DNA linear PAT 20-DEC-2002
 LOCUS AR228552
 DEFINITION Sequence 217 from patent US 648043.
 ACCESSION AR228552
 VERSION AR228552.1 GI:27267389
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1983)
 Chou, G.H., Bailey, C., Hromockyj, A. and Kunsch, C.A.

TITLE Enterococcus faecalis EF040 and uses therefor
JOURNAL Patent: US 6448043-A 217 10-SEP-2002;
Human Genome Sciences, Inc.; Rockville, MD;

FEATURES
source Location/Qualifiers
1..1983
/organism="unknown"

ORIGIN
/mol_type="genomic DNA"

Query Match 100.0%; Score 1881; DB 2; Length 1983;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAAGCATTTAAAAAGTTTGGTACACCGTTAGTACCTTGTACTAATTTTGGCACTT 60
DB |||
QY 100 ATGAAGCATTTAAAAAGTTTGGTACACCGTTAGTACCTTGTACTAATTTTGGCACTT 159
DB |||
QY 61 TTCACAAAGTGTATTAGGACAACAACCTGCACTTTGCAGAGAAAAATGGGAGAGCGACAG 120
DB |||
QY 160 TTCACAAAGTGTATTAGGACAACAACCTGCACTTTGCAGAGAAAAATGGGAGAGCGACAG 219
DB |||
QY 121 CTCGTGATTACAAAAAGAAAATGACGGAATTAACGATCCGCTTAATTCAAAATAGCGGG 180
DB |||
QY 220 CTCGTGATTACAAAAAGAAAATGACGGAATTAACGATCCGCTTAATTCAAAATAGCGGG 279
DB |||
QY 181 AAAAGAAATAGCGAGTTTGAATAATTCAGAGACTGGCAGATGTGACGTTTAGTATTAT 240
DB |||
QY 280 AAAAGAAATAGCGAGTTTGAATAATTCAGAGACTGGCAGATGTGACGTTTAGTATTAT 339
DB |||
QY 241 AACGTGACGAAGAAATTTTACGACGACGAGCGGCAAGCGCAAGCGTTGATGCACTAAA 300
DB |||
QY 340 AACGTGACGAAGAAATTTTACGACGACGAGCGGCAAGCGGCAAGCGTTGATGCACTAAA 399
DB |||
QY 301 CAAGCTGTCGAAGTTTAACTCTCGGGAAAACCTGTGCTCAAGGAACCAACGATGCAAA 360
DB |||
QY 400 CAAGCTGTCGAAGTTTAACTCTCGGGAAAACCTGTGCTCAAGGAACCAACGATGCAAA 459
DB |||
QY 361 GGGAAATGTCAGTTCAGTTTAACTTAAAAACAAAATGTTAAAGTGCATGTATACCAAT 420
DB |||
QY 460 GGGAAATGTCAGTTCAGTTTAACTTAAAAACAAAATGTTAAAGTGCATGTATACCAAT 519
DB |||
QY 421 AAAAGAAACCAAAAAGAGGGTGTAGTGTGCTACGATCAGATGTGGTGGCTTCCAGTT 480
DB |||
QY 520 AAAAGAAACCAAAAAGAGGGTGTAGTGTGCTACGATCAGATGTGGTGGCTTCCAGTT 579
DB |||
QY 481 TAGCAATGATCAAGCAACAGATGTTCTTAAATATATGAAACAGAGAAATTAAGCGGTT 540
DB |||
QY 580 TAGCAATGATCAAGCAACAGATGTTCTTAAATATATGAAACAGAGAAATTAAGCGGTT 639
DB |||
QY 541 GTTCATATTATCTTAAAAATGTTAGGCAATGATGTAATGTAATGTAATGTAATGTAAT 600
DB |||
QY 640 GTTCATATTATCTTAAAAATGTTAGGCAATGATGTAATGTAATGTAATGTAATGTAAT 699
DB |||
QY 601 GGAAGTGTGAAAAATGAAGATTAATGAGCGCAGATTTGTATTTCTAAAAACGAAGGC 660
DB |||
QY 700 GGAAGTGTGAAAAATGAAGATTAATGAGCGCAGATTTGTATTTCTAAAAACGAAGGC 759
DB |||
QY 661 TCACCAAGGCAAGTAAATATATCCAGAGTCAAGATGATTAATATCATGACACAG 720
DB |||
QY 760 TCACCAAGGCAAGTAAATATATCCAGAGTCAAGATGATTAATATCATGACACAG 819
DB |||
QY 721 GATTAAGAAACAGCAAAAGCTTTATTAATGCGGAAAAGTTAATGAAATGGCGAAATGAT 780
DB |||
QY 820 GATTAAGAAACAGCAAAAGCTTTATTAATGCGGAAAAGTTAATGAAATGGCGAAATGAT 879
DB |||
QY 781 TTCACAGAAAGCAGAGATGAAACGAGAGATTAACAGTTAAAAATCTTGAGTTGGTTG 840
DB |||
QY 880 TTCACAGAAAGCAGAGATGAAACGAGAGATTAACAGTTAAAAATCTTGAGTTGGTTG 939
DB |||
QY 841 TATATTTTGAAGAGTAAAGCTCCAAATATATGCAATTAATGAAATCAACAAAA 900
DB |||
QY 940 TATATTTTGAAGAGTAAAGCTCCAAATATATGCAATTAATGAAATCAACAAAA 999
DB |||
```

```
QY 901 AACACATTTACATTTGAAGCAACATCAACACCTGTGTAATGAAAAACGTCAAAATGAT 960
DB |||
QY 1000 AACACATTTACATTTGAAGCAACATCAACACCTGTGTAATGAAAAACGTCAAAATGAT 1059
DB |||
QY 961 ACCCTTAAAGTGTATTAACCAACCAACGCTTGAATGTAAGATGTGGCAATTTGGCGAA 1020
DB |||
QY 1060 ACCCTTAAAGTGTATTAACCAACCAACGCTTGAATGTAAGATGTGGCAATTTGGCGAA 1119
DB |||
QY 1021 AAAATTTAATCAAAATTTCTGTAAATATTCATTTGGGATTCAGACCAAGAGCGAC 1080
DB |||
QY 1120 AAAATTTAATCAAAATTTCTGTAAATATTCATTTGGGATTCAGACCAAGAGCGAC 1179
DB |||
QY 1081 GCTAATTAATAGTCAAAATTTCAATTTAGTGAATTAACATGATGCAAGCCTTAATCTTTGAT 1140
DB |||
QY 1180 GCTAATTAATAGTCAAAATTTCAATTTAGTGAATTAACATGATGCAAGCCTTAATCTTTGAT 1239
DB |||
QY 1141 AACGTGACTTCTGAGAGATATGCTTAATGCGTATATGATGGGATATCAGTGAATGCTCCT 1200
DB |||
QY 1240 AACGTGACTTCTGAGAGATATGCTTAATGCGTATATGATGGGATATCAGTGAATGCTCCT 1299
DB |||
QY 1201 GAAAATTAATCAAGTGACTGAACAGCAATGGCTTCACTGTGCGGTTAATCAGCGTAT 1260
DB |||
QY 1300 GAAAATTAATCAAGTGACTGAACAGCAATGGCTTCACTGTGCGGTTAATCAGCGTAT 1359
DB |||
QY 1261 ATTCTAGGCTTAACAACAGGCGGCAACCTAAATTCGTTTACTTTATGCAATTTAATGAA 1320
DB |||
QY 1360 ATTCTAGGCTTAACAACAGGCGGCAACCTAAATTCGTTTACTTTATGCAATTTAATGAA 1419
DB |||
QY 1321 AAAAGAGATCTTACGAAGAGCTTTAAAAATGAAGGCAATGTTGATTAACGCTATACCGAC 1380
DB |||
QY 1420 AAAAGAGATCTTACGAAGAGCTTTAAAAATGAAGGCAATGTTGATTAACGCTATACCGAC 1479
DB |||
QY 1381 GACCAAAACCAACCAACTGTGAAGTTGTGAACAGTGGGAAAACGTTTCAATTAATGCGAT 1440
DB |||
QY 1480 GACCAAAACCAACCAACTGTGAAGTTGTGAACAGTGGGAAAACGTTTCAATTAATGCGAT 1539
DB |||
QY 1441 GGGAGATGACAGCGCAACCAACCTTTGGCGGAGCTTCTTTGCGTCCGTATCAAAAC 1500
DB |||
QY 1540 GGGAGATGACAGCGCAACCAACCTTTGGCGGAGCTTCTTTGCGTCCGTATCAAAAC 1599
DB |||
QY 1501 AGCGACACAGCAAAATTAATTTGAAAAATCGATGAAACAAACGAAAGCAACCTTGGGTGAAA 1560
DB |||
QY 1600 AGCGACACAGCAAAATTAATTTGAAAAATCGATGAAACAAACGAAAGCAACCTTGGGTGAAA 1659
DB |||
QY 1561 ACAAAGCTGGAAGCAACTTCTTTTCAACAAGCGCTGATGAAATTTGTTGATATCAACAGG 1620
DB |||
QY 1660 ACAAAGCTGGAAGCAACTTCTTTTCAACAAGCGCTGATGAAATTTGTTGATATCAACAGG 1719
DB |||
QY 1621 CTTAAATACGTAACCTATTAATTTAGAGAAACCTGAGCTCGATGATTAATGCTTTGTTA 1680
DB |||
QY 1720 CTTAAATACGTAACCTATTAATTTAGAGAAACCTGAGCTCGATGATTAATGCTTTGTTA 1779
DB |||
QY 1681 ACAAATCGGATTTGAATTTGTGTCAATGAAACATCATATATGCAACAGAAAACTTAAT 1740
DB |||
QY 1780 ACAAATCGGATTTGAATTTGTGTCAATGAAACATCATATATGCAACAGAAAACTTAAT 1839
DB |||
QY 1741 TCACCAAGAAAGTACCAAAACCAACCAAGGTAACCTTACCTTCAACAGTGGCAAAAGGA 1800
DB |||
QY 1840 TCACCAAGAAAGTACCAAAACCAACCAAGGTAACCTTACCTTCAACAGTGGCAAAAGGA 1899
DB |||
QY 1801 ATCTACGTTTACTTGAAGAGTGGCGCAGCTTGCTACTTAATGACAGAGTCTACTTTGCT 1860
DB |||
QY 1900 ATCTACGTTTACTTGAAGAGTGGCGCAGCTTGCTACTTAATGACAGAGTCTACTTTGCT 1959
DB |||
QY 1861 AGACGTTAGAAAAAATAATGCT 1881
DB |||
QY 1960 AGACGTTAGAAAAAATAATGCT 1980
DB |||
```

RESULT 4
AR692967
LOCUS AR692967 1983 bp DNA linear PAT 14-SEP-2005

DEFINITION Sequence 217 from patent US 6913907.
ACCESSION AR692967
VERSION AR692967.1 GI:75182710
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1983)
AUTHORS Choi, G.H., Bailey, C., Hromocky, A. and Kunsch, C.A.
TITLE Enterococcus faecalis polynucleotides encoding EF059
JOURNAL Patent: US 6913907-A 217 05-JUL-2005;
Human Genome Sciences, Inc.; Rockville, MD;
MOX;
FEATURES
source Location/Qualifiers
1..1983
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1881; DB 2; Length 1983;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGCAATTAATAAAGTTGGTACACCGTAGTACCTGTTACTAATTTTGGCCATT 60
DB 100 ATGAAGCAATTAATAAAGTTGGTACACCGTAGTACCTGTTACTAATTTTGGCCATT 159
QY 61 TTCACAGTGTATTAGGGAACAACAACCTGATTTGCAAGAAATAAGGAGAGCGACAG 120
DB 160 TTCACAGTGTATTAGGGAACAACAACCTGATTTGCAAGAAATAAGGAGAGCGACAG 219
QY 121 CTCGTGATTCACAAAAGAAATGACGCGATTACCGATCCGCTTATTCAAAATAGCGGG 180
DB 220 CTCGTGATTCACAAAAGAAATGACGCGATTACCGATCCGCTTATTCAAAATAGCGGG 279
QY 181 AAAAGAAATGAGCGAGTTGTATTAATATCAAGGACTGGCAGATGTGACGTTTATTTAT 240
DB 280 AAAAGAAATGAGCGAGTTGTATTAATATCAAGGACTGGCAGATGTGACGTTTATTTAT 339
QY 241 AACGTGACGAACGAATTTTACGACCAACGAGCGGCGCAAGCGTTGATGACGTAA 300
DB 340 AACGTGACGAACGAATTTTACGACCAACGAGCGGCGCAAGCGTTGATGACGTAA 399
QY 301 CAAAGCTGTCCAAAGTTAATCTCGGGAACCTGTGGTCAAGAAACCAACCGATGCAAT 360
DB 400 CAAAGCTGTCCAAAGTTAATCTCGGGAACCTGTGGTCAAGAAACCAACCGATGCAAT 459
QY 361 GGGAAATGCTACGTTCAGTTACCTAATAAACAATAATGTAAGATGCAAGTATACATT 420
DB 460 GGGAAATGCTACGTTCAGTTACCTAATAAACAATAATGTAAGATGCAAGTATACATT 519
QY 421 AAAAGAAACCAAAAGAGGGTGTAGTTGCTGCTACGAATATGTGTGCGTTCCCAATT 480
DB 520 AAAAGAAACCAAAAGAGGGTGTAGTTGCTGCTACGAATATGTGTGCGTTCCCAATT 579
QY 481 TAGCAAAATGATCAAGCAACAGATGTTCTTATTAATATGSAACAGAAATTAAGGGTT 540
DB 580 TAGCAAAATGATCAAGCAACAGATGTTCTTATTAATATGSAACAGAAATTAAGGGTT 639
QY 541 GTTCAATATTTATCTTAATAATGTGTAGCAATGATGATGTTACATGTGAAAAAGTA 600
DB 640 GTTCAATATTTATCTTAATAATGTGTAGCAATGATGATGTTACATGTGAAAAAGTA 699
QY 601 GGAAGCTGTGAATAAGAGATTAAATGGCGCAGAATTTGTTATTTCTAAAAAGCGAAGGC 660
DB 700 GGAAGCTGTGAATAAGAGATTAAATGGCGCAGAATTTGTTATTTCTAAAAAGCGAAGGC 759
QY 661 TCACCAAGGACAGTAATAATATATCAAGAGTCAAGATGATATATATCATGSAACAAG 720
DB 760 TCACCAAGGACAGTAATAATATATCAAGAGTCAAGATGATATATATCATGSAACAAG 819
QY 721 GATTAAGAAACAAGCAAAACGCTTTATTAATGGAAGAAATGTAATGCAATGGCAAAATGAT 780

DB 820 GATTAAGAAACAAGCAAAACGCTTTATTAATGGAAGAAATGTAATGCAATGGCAAAATGAT 879
QY 781 TTCACGAAGCAGAGATGAAACGGAGAAATTAACGTTAATAATCTTGAAGTTGGTTGG 840
DB 880 TTCACGAAGCAGAGATGAAACGGAGAAATTAACGTTAATAATCTTGAAGTTGGTTGG 939
QY 841 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAGATTAATGAAATCAAAACAAA 900
DB 940 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAGATTAATGAAATCAAAACAAA 999
QY 901 ACACCATTTTACATTTGAAACCAACATCAACACCTGTGAAAAAAACATCAAAATGAT 960
DB 1000 ACACCATTTTACATTTGAAACCAACATCAACACCTGTGAAAAAAACATCAAAATGAT 1059
QY 961 ACCTCTAAAGTTGATTAACCAACACCAAGCTTATGATGTTAAGATGTGCAATGGCGAA 1020
DB 1060 ACCTCTAAAGTTGATTAACCAACACCAAGCTTATGATGTTAAGATGTGCAATGGCGAA 1119
QY 1021 AAAATTTAATTAATCAAAATTTCTGTAAATATTCATTTGGGGATTTGACACAAAGGCGAC 1080
DB 1120 AAAATTTAATTAATCAAAATTTCTGTAAATATTCATTTGGGGATTTGACACAAAGGCGAC 1179
QY 1081 GCTAATTAATTAATGCTCAAAATTTCAATTTAGTTGAATAACATGATGACGCTTAACTTTGAT 1140
DB 1180 GCTAATTAATTAATGCTCAAAATTTCAATTTAGTTGAATAACATGATGACGCTTAACTTTGAT 1239
QY 1141 AACGTGACTTCTGAGAGATATGCTTATGCGTTATATGATGAGGAGATCAAGTATGCTCCT 1200
DB 1240 AACGTGACTTCTGAGAGATATGCTTATGCGTTATATGATGAGGAGATCAAGTATGCTCCT 1299
QY 1201 GAAATTTATCAAGTATGCTGAACAAGCAATGCGTTCACTGTGGCGCTTAATCCAGGTTAT 1260
DB 1300 GAAATTTATCAAGTATGCTGAACAAGCAATGCGTTCACTGTGGCGCTTAATCCAGGTTAT 1359
QY 1261 ATTCTACGCTTAACCAACGCGGCACTAATAATTCGTTTACTTTATGCAATTAATGA 1320
DB 1360 ATTCTACGCTTAACCAACGCGGCACTAATAATTCGTTTACTTTATGCAATTAATGA 1419
QY 1321 AAAGCAATCTTAACGAAGGCTTTAATAATGAGCGAATGTTGATACGCTACACGAC 1380
DB 1420 AAAGCAATCTTAACGAAGGCTTTAATAATGAGCGAATGTTGATACGCTACACGAC 1479
QY 1381 GACCAAAACCAACCAACGTTGAAGTTGACAGGGGGGAAACGTTTCAATTAAGTCAT 1440
DB 1480 GACCAAAACCAACCAACGTTGAAGTTGACAGGGGGGAAACGTTTCAATTAAGTCAT 1539
QY 1441 GCGGATGTGACAGCAACAACGCTTGGCGGAGCTTCTTGTGCTCCGTATCAAAAC 1500
DB 1540 GCGGATGTGACAGCAACAACGCTTGGCGGAGCTTCTTGTGCTCCGTATCAAAAC 1599
QY 1501 AGCGACACGCAAAATTTTGAATAATCGATGAACAAGAAACGAACTTGGGTGAAA 1560
DB 1600 AGCGACACGCAAAATTTTGAATAATCGATGAACAAGAAACGAACTTGGGTGAAA 1659
QY 1561 ACAAAAGCTGAAGCAACTTCTTCAACAAGGCGATGATTTGTTGATATCAACAGGG 1620
DB 1660 ACAAAAGCTGAAGCAACTTCTTCAACAAGGCGATGATTTGTTGATATCAACAGGG 1719
QY 1621 CTTAATAATCGTACCTAATTTATTAAGAAACCTGATGCTCTGATGATTAATGCTTGT 1680
DB 1720 CTTAATAATCGTACCTAATTTATTAAGAAACCTGATGCTCTGATGATTAATGCTTGT 1779
QY 1681 ACAAAATCGGATTTGATTTGTGCTAATGAACAATATGCAACAAGAAACCTAAGTT 1740
DB 1780 ACAAAATCGGATTTGATTTGTGCTAATGAACAATATGCAACAAGAAACCTAAGTT 1839
QY 1741 TCACCAAGAAAGTACCAAAACAACAAGGTAACCTTACCTTCAACAGTGGCAAGGA 1800
DB 1840 TCACCAAGAAAGTACCAAAACAACAAGGTAACCTTACCTTCAACAGTGGCAAGGA 1899
QY 1801 ATCTAGTTTACTTATGAAAGTGGCGAGCTTGTACTTATTTGACAGAGTCTAATTTGCT 1860
DB 1900 ATCTAGTTTACTTATGAAAGTGGCGAGCTTGTACTTATTTGACAGAGTCTAATTTGCT 1959

QY 1861 AGACGTAGAAAAGAAATGCT 1881
Db 1960 AGACGTAGAAAAGAAATGCT 1980

RESULT 5
BD193451

LOCUS BD193451 15614 bp DNA linear PAT 17-UTL-2003

DEFINITION Enterococcus faecialis polynucleotides and polypeptides.

ACCESSION BD193451

VERSION BD193451.1 GI:33003190

KEYWORDS JP 2002529046-A/45.

SOURCE unidentified

ORGANISM unidentified

REFERENCE 1 (bases 1 to 15614)
Kunsch,C.A., Dillon,P.J. and Baraeh,S.C.
Enterococcus faecialis polynucleotides and polypeptides
Parent: JP 2002529046-A 45 03-SEP-2002;
JOURNAL HUMAN GENOME SCIENCES INC

COMMENT OS Unidentified
PN JP 2002529046-A/45
PD 03-SEP-2002
PF 04-MAY-1998 JP 1998548302
PR 06-MAY-1997 US 60/044031,16-MAY-1997 US 60/046655 PR
14-NOV-1997 US 60/066009
PI CHARLES A KUNSCH,PATRICK J DILLON,STEVEN C BARASH PC
CI2N15/31,C07K14/315,C07K16/12,C1201/68
CC Strandedness: Double;
CC Topology: Linear;
CC Enterococcus faecialis polynucleotides and polypeptides FH
Key Location/Qualifiers
FT source 1..15614

FEATURES
source 1..15614
Location/Qualifiers
1..15614
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 1881; DB 2; Length 15614;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCAATTAAAAAGTTGGTACACCGTTAGTACCTGTTACTAATTTTGCACATT 60
Db 12164 ATGAAGCAATTAAAAAGTTGGTACACCGTTAGTACCTGTTACTAATTTTGCACATT 12223

QY 61 TTCACAAGGTATTAGGACCAACACTGATTTCAGAGAATAATGGGAGAGCGACAG 120
Db 12224 TTCACAAGGTATTAGGACCAACACTGATTTCAGAGAATAATGGGAGAGCGACAG 12283

QY 121 CTGCGATTTCACAAAAAGAAATGACGATTTTACGATCCGCTTATTCAAAAATAGCGGG 180
Db 12284 CTGCGATTTCACAAAAAGAAATGACGATTTTACGATCCGCTTATTCAAAAATAGCGGG 12343

QY 181 AAAAGAAATGAGGAGTTGATTAATATCAAGACTGCGATGATGACGTTTATGATTTAT 240
Db 12344 AAAAGAAATGAGGAGTTGATTAATATCAAGACTGCGATGATGACGTTTATGATTTAT 12403

QY 241 AACGTGACGAAGCAATTTTACGAGCAACGAGCGGACGAGGACGTTGATGACGCTAATA 300
Db 12404 AACGTGACGAAGCAATTTTACGAGCAACGAGCGGACGAGGACGTTGATGACGCTAATA 12463

QY 301 CAAGCTGTCCAAAGTTTAACTCTCGGAAAACCTGTTGCTCAAGGAACCAACGATGCAAT 360
Db 12464 CAAGCTGTCCAAAGTTTAACTCTCGGAAAACCTGTTGCTCAAGGAACCAACGATGCAAT 12523

QY 361 GGGAAATGTCAGTGTAGTTACTAATAAAACAAAATGATGAAGATGATGATTAACATT 420
Db 12524 GGGAAATGTCAGTGTAGTTACTAATAAAACAAAATGATGAAGATGATGATTAACATT 12583

QY 421 AAAAGAAACCAAAAGAGGTGTAGTTGCTGTACGAATATGTTGGGCGTTCCAGTT 480
Db 12584 AAAAGAAACCAAAAGAGGTGTAGTTGCTGTACGAATATGTTGGGCGTTCCAGTT 12643

QY 481 TACGAATGATCAACCAACAGATGTTCTTAAATATGGAACAGAAATTTAGCGGTT 540
Db 12644 TACGAATGATCAACCAACAGATGTTCTTAAATATGGAACAGAAATTTAGCGGTT 12703

QY 541 GTTCAATTTTATCTTAAAAATGTGTAGCCATGATGTAAGTTTCAATGTAAAAAGTA 600
Db 12704 GTTCAATTTTATCTTAAAAATGTGTAGCCATGATGTAAGTTTCAATGTAAAAAGTA 12763

QY 601 GGAACGTGCAAAAATGAAAGATTAATATGCGGAGATTTGTTATTTCTTAAAAAGCAAGC 660
Db 12764 GGAACGTGCAAAAATGAAAGATTAATATGCGGAGATTTGTTATTTCTTAAAAAGCAAGC 12823

QY 661 TCACGAGGCAAGTAAATATATCCAGAGAGTCAAAAGATGATTAATATCATGACAAAG 720
Db 12824 TCACGAGGCAAGTAAATATATCCAGAGAGTCAAAAGATGATTAATATCATGACAAAG 12883

QY 721 GATTAAGAAACAAGCAAAACGTTTATTACTGGAAGATTAAGTAATGGCGAAATGAT 780
Db 12884 GATTAAGAAACAAGCAAAACGTTTATTACTGGAAGATTAAGTAATGGCGAAATGAT 12943

QY 781 TTCACAGAGCAGAGAAATGAAACGGGAGAAATTAACATTAATAATCTTGAGGTTGTTG 840
Db 12944 TTCACAGAGCAGAGAAATGAAACGGGAGAAATTAACATTAATAATCTTGAGGTTGTTG 13003

QY 841 TATATTTTGAAGAGTAAAGTCCCAATTAATGCAAGATTAATGAAATCAACAAA 900
Db 13004 TATATTTTGAAGAGTAAAGTCCCAATTAATGCAAGATTAATGAAATCAACAAA 13063

QY 901 ACACATTTTACATTTGAAGCAAAACATCAACACCTGTTGAAAAACGTCAAAATGAT 960
Db 13064 ACACATTTTACATTTGAAGCAAAACATCAACACCTGTTGAAAAACGTCAAAATGAT 13123

QY 961 ACCCTAAAGTTGATTAACCAACACCAAGCTTAGATGTAAGATGTCGCAATTTGGCANA 1020
Db 13124 ACCCTAAAGTTGATTAACCAACACCAAGCTTAGATGTAAGATGTCGCAATTTGGCANA 13183

QY 1021 AAAATTTAATCAAAATTTCTGTAAATATTCATTGGGGATTGCAACAAAGAGCGAC 1080
Db 13184 AAAATTTAATCAAAATTTCTGTAAATATTCATTGGGGATTGCAACAAAGAGCGAC 13243

QY 1081 GCTAATTAATCGTCAAAATTCATTAATGATTAACATGATGACGCTTAACTTTGAT 1140
Db 13244 GCTAATTAATCGTCAAAATTCATTAATGATTAACATGATGACGCTTAACTTTGAT 13303

QY 1141 AACGTACTTCTGAGAGATGCTTATGCGTTATATGATGGGATACAGTGAATGCTCT 1200
Db 13304 AACGTACTTCTGAGAGATGCTTATGCGTTATATGATGGGATACAGTGAATGCTCT 13363

QY 1201 GAAATTTATCAAGTACGTAACACCAATGCTTCACTGTGCGGCTTAACTCAACGAT 1260
Db 13364 GAAATTTATCAAGTACGTAACACCAATGCTTCACTGTGCGGCTTAACTCAACGAT 13423

QY 1261 ATTCTAGGCTTAACACGAGCGGCACTAAATTCGTTTACTTTATGCAATTAATGTA 1320
Db 13424 ATTCTAGGCTTAACACGAGCGGCACTAAATTCGTTTACTTTATGCAATTAATGTA 13483

QY 1321 AAAGCAGATCTTACGAAAGGCTTTAAAAATGAGGCAATGTTGATPAACGGTCATACGAC 1380
Db 13484 AAAGCAGATCTTACGAAAGGCTTTAAAAATGAGGCAATGTTGATPAACGGTCATACGAC 13543

QY 1381 GACCAAAACCAACCAACTGTTGAAGTTGTAAGGTGGGAAACGTTTCAATTAATAGTCGAT 1440
Db 13544 GACCAAAACCAACCAACTGTTGAAGTTGTAAGGTGGGAAACGTTTCAATTAATAGTCGAT 13603

QY 1441 GCGAGATGACAGGACCAACGCTTGGCGGAGCTTCTTTGTGTCGTCGATCAAAAC 1500
Db 13604 GCGAGATGACAGGACCAACGCTTGGCGGAGCTTCTTTGTGTCGTCGATCAAAAC 13663

Qy	1501	AGCGCACGAGCAATTTATTTGAAATGATGAAACAACGAAAGCGCAACTTGGGTGAAA	1560
Db	13664	AGCGCACGAGCAATTTATTTGAAATGATGAAACAACGAAAGCGCAACTTGGGTGAAA	13723
Qy	1561	ACAAAGGTGAGCACTACTTTTACAAACAACGGCTGATGATTTAGTATATACACGGG	1620
Db	13724	ACAAAGGTGAGCACTACTTTTACAAACAACGGCTGATGATTTAGTATATACACGGG	13783
Qy	1621	CTTAAATACGGTACCTATTTATTTAGAAACACTGAGCTTCCTGATGATTAATGCTTGTTA	1680
Db	13784	CTTAAATACGGTACCTATTTATTTAGAAACACTGAGCTTCCTGATGATTAATGCTTGTTA	13843
Qy	1681	ACAAATGCGATTTGAATTTGTGTCAATGAAACAATCATATGCGCAACAAGAAAACCTTAGTT	1740
Db	13844	ACAAATGCGATTTGAATTTGTGTCAATGAAACAATCATATGCGCAACAAGAAAACCTTAGTT	13903
Qy	1741	TCACCAGAAAAAGTACCAAAACAACAAGATACCTTACCTTCAACAGGTGCAAAAGGA	1800
Db	13904	TCACCAGAAAAAGTACCAAAACAACAAGATACCTTACCTTCAACAGGTGCAAAAGGA	13963
Qy	1801	ATCTACGTTTACTTAGAGTGTGGCGCACTTGTCTACTTATTTGACAGAGCTTACTTTGCT	1860
Db	13964	ATCTACGTTTACTTAGAGTGTGGCGCACTTGTCTACTTATTTGACAGAGCTTACTTTGCT	14023
Qy	1861	AGACGTGAAAAAGAAAATGCT 1881	
Db	14024	AGACGTGAAAAAGAAAATGCT 14044	

[illegible]

ORGANISM	SOURCE
<i>Enterococcus faecalis</i>	faecal
<i>Enterococcus faecalis</i>	faecal

REFERENCE AUTHORS	
1	Meinke, A., Nagy, E., Hanner, M. and Gelbmann, D.

Patent: WO 2004106367-A 28 09-DEC-2004;
JOURNAL

INTERCELL AG (A1)
Location/Qualifiers

```

/organism="Enterococcus faecalis"
/mol_type="unassigned DNA"
/db_xref="caxon:1351"

```

ORIGIN

Query Match	98.8%	Score 1859;	DB 2;	Length 1875;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches 1875; Conservative	0;	Mismatches	6;	Gaps 1;

Oy		1	ATGAAGCAATTAAAAAAAGTTTGGTAACACCCTTAGTAACCTTGTTACTAATTTTGCACATT	60
Db		1	ATGAAGCAATTAAAAAAAGTTTGGTAACACCCTTAGTAACCTTGTTACTAATTTTGCACATT	60
Oy		61	TTCAACAAGTATTTAGGGAACAACACTGCATTTCAGAAGAAAATGGGGAGAGCGCACAG	120
Db		61	TTCAACAAGTATTTAGGGAACAACACTGCATTTCAGAGAAAATGGGGAGAGCGCACAG	120
Oy		121	CTCGGATTCACAAAAGAAAAATGACGGAATTTACAGATCCGCTTATTTCAAATATAGCGGG	180
Db		121	CTCGGATTCACAAAAGAAAAATGACGGAATTTACAGATCCGCTTATTTCAAATATAGCGGG	180
Oy		181	AAAGAAATAGAGCGAATTTGATAAATATCAAGACCTGGCAGATGTGACGTTTTAGTATTTAT	240
Db		181	AAAGAAATAGAGCGAATTTGATAAATATCAAGACCTGGCAGATGTGACGTTTTAGTATTTAT	240
Oy		241	AACGTGACGAACGAATTTTACAGCAACGAGCGGCAGGCGCAAGCGTTGATACAGCTAA	300
Db		241	AACGTGACGAACGAATTTTACAGCAACGAGCGGCAGGCGCAAGCGTTGATACAGCTAA	300

Db	241	AACTGAGAAAGAAATTTTACGACCAAGACGGACGCAAGGCTTGAATGCACTAA	300
OY	301	CAAGCTGCCAAAGTTTAACTCCTGGGAAAACCTGTTGCTCAAGAAACCAACGATCAAT	360
Db	301	CAGCTGCCAAAGTTTAACTCCTGGGAAAACCTGTTGCTCAAGAAACCAACGATCAAT	360
OY	361	GGAATGTCACGTGTTCACTTACCTTAAAAAAMAAATGCTAAAGATGCACTATACAT	420
Db	361	GGAATGTCACGTGTTCACTTACCTTAAAAAAMAAATGCTAAAGATGCACTATACAT	420
OY	421	AAAGAAAGCAAAAGAGGCTGTAGTGTCTGCTACGAATATGCTGTGGCTTCCCACTT	480
Db	421	AAAGAAAGCAAAAGAGGCTGTAGTGTCTGCTACGAATATGCTGTGGCTTCCCACTT	480
OY	481	TACGAAATGATCAAGCAACAGATGGTTCCTATPAAATATGGAACAGAAATTTGCGGTT	540
Db	481	TACGAAATGATCAAGCAACAGATGGTTCCTATPAAATATGGAACAGAAATTTGCGGTT	540
OY	541	GTTCAATATTTATCCTAAAAATGTGTAGCCAAATGCTAGCTTTACATGTAAAAAGTA	600
Db	541	GTTCAATATTTATCCTAAAAATGTGTAGCCAAATGCTAGCTTTACATGTAAAAAGTA	600
OY	601	GGAACCTGCTAAAAATGAAGATTAAATGGCCAGATTTGTTATTTCTPAAAAGCGAAGC	660
Db	595	GGAACCTGCTAAAAATGAAGATTAAATGGCCAGATTTGTTATTTCTPAAAAGCGAAGC	654
OY	661	TCACAGGCAACGTPAAAATATATCCAGGAATCCAAAGATGAATATATATACATGCAACG	720
Db	655	TCACAGGCAACGTPAAAATATATCCAGGAATCCAAAGATGAATATATATACATGCAACG	714
OY	721	GATAAAGCAACAGCAAAAACGCTTTATTTACTGGAAAAAGTTATGAAATTTGGCAAAATGAT	780
Db	715	GATAAAGCAACAGCAAAAACGCTTTATTTACTGGAAAAAGTTATGAAATTTGGCAAAATGAT	774
OY	781	TTCAACAGACGAGATGTAACGGAGAAATTAACATGTTAAAAATCTTGAGGTTGGTTCC	840
Db	775	TTCAACAGACGAGATGTAACGGAGAAATTAACATGTTAAAAATCTTGAGGTTGGTTCC	834
OY	841	TATATTTTAAAGAAAGTAAAAGCTCCAAATTAATGCAAGATTAATTTGAAAAATCAACAAA	900
Db	835	TATATTTTAAAGAAAGTAAAAGCTCCAAATTAATGCAAGATTAATTTGAAAAATCAACAAA	894
OY	901	ACACCAATTTCAATTTGAAGCAAAACATCAAAACACCTGTTGAAAAAACAAGTCAAAAATGAT	960
Db	895	ACACCAATTTCAATTTGAAGCAAAACATCAAAACACCTGTTGAAAAAACAAGTCAAAAATGAT	954
OY	961	ACCTCTAAAGTTGATAAAACAACAACCAAGCTTAGATGCTAAAGATGTGGCAATTTGGCGAA	1020
Db	955	ACCTCTAAAGTTGATAAAACAACAACAACCAAGCTTAGATGCTAAAGATGTGGCAATTTGGCGAA	1014
OY	1021	AAAAATTAATATCAAAATTTCTGTAAATATTCATTTGGGAAATTCGACACAAAGAAAGCGAC	1080
Db	1015	AAAAATTAATATCAAAATTTCTGTAAATATTCATTTGGGAAATTCGACACAAAGAAAGCGAC	1074
OY	1081	GCTATATAATACGTCAAAATTTGTTGTTGATATAACATGATGCAAGCTTTAACTTTTGAT	1140
Db	1075	GCTATATAATACGTCAAAATTTGTTGTTGATATAACATGATGCAAGCTTTAACTTTTGAT	1134
OY	1141	AACGTGACTTGGAGAGATATGCTTATGCGTTATATGATGAGGAAATACAGTATTTGCTCT	1200
Db	1135	AACGTGACTTGGAGAGATATGCTTATGCGTTATATGATGAGGAAATACAGTATTTGCTCT	1194
OY	1201	GAAAAATATCAAGATGATGTAACAAGCAAAATGGCTTCACTGTGCGCGCTTAAATCCAGCGAT	1266
Db	1195	GAAAAATATCAAGATGATGTAACAAGCAAAATGGCTTCACTGTGCGCGCTTAAATCCAGCGAT	1254
OY	1261	ATTCTCAAGCTAACACCAAGGCGGCACTPAAAAATTCGTTTACTTATGCAATTTAAATGAA	1320
Db	1255	ATTCTCAAGCTAACACCAAGGCGGCACTPAAAAATTCGTTTACTTATGCAATTTAAATGAA	1314
OY	1321	AAACGAGATCTACGAAAGGCTTTAAAAATGAGCGAATGTTGATPAAACGTCATACCGAC	1380

Qy	1075	GCCTAATAATACGTGCAAAATTCGAAATTTAGTGTATTAACATGATGCAAGCCTTAACTTTGAT	1134
Db	1075	GCCTAATAATACGTGCAAAATTCGAAATTTAGTGTATTAACATGATGCAAGCCTTAACTTTGAT	1134
Qy	1141	AACGTGACTTCGAGAGATATGCTTATGCGTTAATGATGGGATACAGTATTCCTCT	1200
Db	1135	AACGTGACTTCGAGAGATATGCTTATGCGTTAATGATGGGATACAGTATTCCTCT	1194
Qy	1201	GAAATTTATCAGTGACTGAACAACAGCAATAGGCTTCACTGTGCGCGTTAATCCAGCGAT	1260
Db	1195	GAAATTTATCAGTGACTGAACAACAGCAATAGGCTTCACTGTGCGCGTTAATCCAGCGAT	1254
Qy	1261	ATTCCCTACGTACACCAAGCGGCGACACTAAATTCGTTACTTTATGCAATTTAATGAA	1320
Db	1255	ATTCCCTACGTACACCAAGCGGCGACACTAAATTCGTTACTTTATGCAATTTAATGAA	1314
Qy	1321	AAAGCAGATCCTACGAAAGGCTTTAAATAGAGCGAATGTTGATTAACGATACCGAC	1380

Db 1315 AAAGCAGATCTACGAAAGCTTTAAATAATGAGCGAATGTTGATTAAGCTATACCGAC 1374
Qy 1381 GACCAAAACCAACCACTGTTGAAGTTTGACAGGTGGGAAACGTTTCTTAAAGTCAT 1440
Db 1375 GACCAAAACCAACCACTGTTGAAGTTTGACAGGTGGGAAACGTTTCTTAAAGTCAT 1434
Qy 1441 GCGCATGTGACGACGACACAGCCTTGGGGAGCTTCTTTGCGTCGATCAAAAC 1500
Db 1435 GCGCATGTGACGACGACACAGCCTTGGGGAGCTTCTTTGCGTCGATCAAAAC 1494
Qy 1501 AGCGACACGCAAAATTTATTTGAAATCGATGAAACAGAAACAGCAACTTGGTGAA 1560
Db 1495 AGCGACACGCAAAATTTATTTGAAATCGATGAAACAGAAACAGCAACTTGGTGAA 1554
Qy 1561 ACAAAGCTGAAGCACTCTTTTACAAACGGCTGATGATTAAGTTATATCAGAG 1620
Db 1555 ACAAAGCTGAAGCACTCTTTTACAAACGGCTGATGATTAAGTTATATCAGAG 1614
Qy 1621 CTTAATACGGTACTATTTATTTAGAAAGCTGTAGCTCGATGATTAAGTTATGTT 1680
Db 1615 CTTAATACGGTACTATTTATTTAGAAAGCTGTAGCTCGATGATTAAGTTATGTT 1674
Qy 1681 ACAAATCGGATTTGAAATTTGTGTCAATGAAACAATCATATGACCAACAGAAACCTAGT 1740
Db 1675 ACAAATCGGATTTGAAATTTGTGTCAATGAAACAATCATATGACCAACAGAAACCTAGT 1734
Qy 1741 TCACAGAAAGAAAGTACCAAGCAACAAAGTACCTTCTTCAACAGGTGGCAAGAA 1800
Db 1735 TCACAGAAAGAAAGTACCAAGCAACAAAGTACCTTCTTCAACAGGTGGCAAGAA 1794
Qy 1801 ATCTAGTTTACTTGAAGTGGCGCAGTCTTGTACTTATTTGAGAGTCTTACTTGT 1860
Db 1795 ATCTAGTTTACTTGAAGTGGCGCAGTCTTGTACTTATTTGAGAGTCTTACTTGT 1854
Qy 1861 AGACGTAGAAAGAAATGCT 1881
Db 1855 AGACGTAGAAAGAAATGCT 1875

RESULT 7
AE016830_10
WPCOMMENT

Sequence split into 33 fragments LOCUS AE016830 Accession AE016830
Fragment Name Begin End
AE016830_01 1 110000
AE016830_02 20001 310000
AE016830_03 30001 410000
AE016830_04 40001 510000
AE016830_05 50001 610000
AE016830_06 60001 710000
AE016830_07 70001 810000
AE016830_08 80001 910000
AE016830_09 90001 1010000
AE016830_10 100001 1110000
AE016830_11 110001 1210000
AE016830_12 120001 1310000
AE016830_13 130001 1410000
AE016830_14 140001 1510000
AE016830_15 150001 1610000
AE016830_16 160001 1710000
AE016830_17 170001 1810000
AE016830_18 180001 1910000
AE016830_19 190001 2010000
AE016830_20 200001 2110000
AE016830_21 210001 2210000
AE016830_22 220001 2310000
AE016830_23 230001 2410000
AE016830_24 240001 2510000
AE016830_25 250001 2610000
AE016830_26 260001 2710000
AE016830_27 270001 2810000
AE016830_28 280001 2910000

AE016830_29 290001 3010000
AE016830_30 300001 3110000
AE016830_31 310001 3210000
AE016830_32 320001 3218031
Continuation (11 of 33) of AE016830 from base 1000001 (AE016830 Enterococcus faecalis V5
Query Match 98.8%; Score 1859; DB 15; Length 110000;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1875; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
Qy 1 ATGAAGCAATTAATAAAAGTTTGTGACACCGTTAGTCTGTATCTTAATTTTGCACCT 60
Db 62129 ATGAAGCAATTAATAAAAGTTTGTGACACCGTTAGTCTGTATCTTAATTTTGCACCT 62188
Qy 61 TTCACAAAGTATTTGGGACCAACCTGATTTTGGAAAGAAAGGGGAGCGCACAG 120
Db 62189 TTCACAAAGTATTTGGGACCAACCTGATTTTGGAAAGAAAGGGGAGCGCACAG 62248
Qy 121 CTGATTTTCAAAAAGAAATGACGATTTTACAGATCCGCTTATTTCAAAATAGCGG 180
Db 62249 CTGATTTTCAAAAAGAAATGACGATTTTACAGATCCGCTTATTTCAAAATAGCGG 62308
Qy 181 AAAGAAATGACGAGTTTGTATTAATATCAAGACTGGCAGATGTGACGTTTAGTAT 240
Db 62309 AAAGAAATGACGAGTTTGTATTAATATCAAGACTGGCAGATGTGACGTTTAGTAT 62368
Qy 241 AACGTGACGAAGATTTTATGAGACAGAGCGGACGAGGTTGATGACGCTAA 300
Db 62369 AACGTGACGAAGATTTTATGAGACAGAGCGGACGAGGTTGATGACGCTAA 62428
Qy 301 CAAGCTGTCAAAGTTTAACTCTCGGAAACCTGTGTCTCAAGAAACCAACGATGCAAT 360
Db 62429 CAAGCTGTCAAAGTTTAACTCTCGGAAACCTGTGTCTCAAGAAACCAACGATGCAAT 62488
Qy 361 GGGAAATGTCACTGTTCAAGTTACTTAAATAAAACAAATGTTAAAGATGCAATACATT 420
Db 62489 GGGAAATGTCACTGTTCAAGTTACTTAAATAAAACAAATGTTAAAGATGCAATACATT 62548
Qy 421 AAAGAAAGACCAAAAGAGGTGTAGTGTCTGCTAGCAATATGATGAGCGTCCCAAGT 480
Db 62549 AAAGAAAGACCAAAAGAGGTGTAGTGTCTGCTAGCAATATGATGAGCGTCCCAAGT 62608
Qy 481 TAGAAATGATCAACAAACAGATGTTCTTAATAATGAAACAGAAATTAAGCGTT 540
Db 62609 TAGAAATGATCAACAAACAGATGTTCTTAATAATGAAACAGAAATTAAGCGTT 62668
Qy 541 GTTCATATTTATCTTAATAATGTGTAGCCATGATGATGTTTATGATGAAAAAGTA 600
Db 62669 GTTCATATTTATCTTAATAATGTGTAGCCATGATGATGTTTATGATGAAAAAGTA 62722
Qy 601 GGAATGCTGAATAATGAAGATTAATATGCGCAGAAATTTGTTATTTCTAAAGCAAGC 660
Db 62723 GGAATGCTGAATAATGAAGATTAATATGCGCAGAAATTTGTTATTTCTAAAGCAAGC 62782
Qy 661 TCACAGGCAAGTAAATATATCCAGAGATCAAGATGATTAATATATCATGACACAG 720
Db 62783 TCACAGGCAAGTAAATATATCCAGAGATCAAGATGATTAATATATCATGACACAG 62842
Qy 721 GATTAAGAACAGCAAAACGCTTTATTACTGGGAAAAAGTTAAGAAATGGCGAAATGAT 780
Db 62843 GATTAAGAACAGCAAAACGCTTTATTACTGGGAAAAAGTTAAGAAATGGCGAAATGAT 62902
Qy 781 TTCACAGAAAGCAGAAATGAAACGGAGATTAACAGTTTAAATAATCTTGAGTTGGT 840
Db 62903 TTCACAGAAAGCAGAAATGAAACGGAGATTAACAGTTTAAATAATCTTGAGTTGGT 62962
Qy 841 TATATTTTGAAGAGTAAAGCTCCAAATTAATGCAATTAATTAATGAAATCAACAAA 900
Db 62963 TATATTTTGAAGAGTAAAGCTCCAAATTAATGCAATTAATTAATGAAATCAACAAA 63022
Qy 901 AACCAATTTAACAATGAAAGCAACATCAACCTGTTGAAATAACGTCAAAAATGAT 960
Db 63023 AACCAATTTAACAATGAAAGCAACATCAACCTGTTGAAATAACGTCAAAAATGAT 63082

```

QY 961 ACCTGAAAGTGTATTAACCAACCAAGCTTAGATGTAAAGATGTGGCAATTGGCCGA 1020
Db 63083 ACCCTTAAGTGTATTAACCAACCAAGCTTAGATGTAAAGATGTGGCAATTGGCCGA 63142
QY 1021 AAAATTAATATCAAAATTTCTGTAAATATATTCATTTGGGATTCAGACAAAGAGCGAC 1080
Db 63143 AAAATTAATATCAAAATTTCTGTAAATATATTCATTTGGGATTCAGACAAAGAGCGAC 63202
QY 1081 GCTAATTAATACGTCAAAATTCATTTAGTTGATTAACATGATGCAAGCTTTAATCTTTGAT 1140
Db 63203 GCTAATTAATACGTCAAAATTCATTTAGTTGATTAACATGATGCAAGCTTTAATCTTTGAT 63262
QY 1141 AACGAGCTTCTGGAGAGATATCTTATATGATGAGGAGATCACTGATGTCCT 1200
Db 63263 AACGAGCTTCTGGAGAGATATCTTATATGATGAGGAGATCACTGATGTCCT 63322
QY 1201 GAAATTAATCAAGTGTAGTAAAGCAAGCAATGGCTTCACTGCGCCGTTAATCAAGCTAT 1260
Db 63323 GAAATTAATCAAGTGTAGTAAAGCAAGCAATGGCTTCACTGCGCCGTTAATCAAGCTAT 63382
QY 1261 ATTCTAGGCTTAACCAAGCGGCACTAAATTCGTTTACTTTATGCAATTTAAATGAA 1320
Db 63383 ATTCTAGGCTTAACCAAGCGGCACTAAATTCGTTTACTTTATGCAATTTAAATGAA 63442
QY 1321 AAAGAGATCCCTACGAAGGCTTTAAATAATGAGGGAATGTTATPAAAGCTATACCGAC 1380
Db 63443 AAAGAGATCCCTACGAAGGCTTTAAATAATGAGGGAATGTTATPAAAGCTATACCGAC 63502
QY 1381 GACCAAAACCAACCACTGTTGATGTGACAGGTGGGAAAGCTTCAATTAAGTCAT 1440
Db 63503 GACCAAAACCAACCACTGTTGATGTGACAGGTGGGAAAGCTTCAATTAAGTCAT 63562
QY 1441 GGCGATGTGACAGCAACAGCCTTTGGGGAGCTTCTTTGTCGTCGATCAAAAC 1500
Db 63563 GGCGATGTGACAGCAACAGCCTTTGGGGAGCTTCTTTGTCGTCGATCAAAAC 63622
QY 1501 AGCGACAGCAAAATTTATTTGAAATCCATGTAACACGAAAGCACTTGGGTGAA 1560
Db 63623 AGCGACAGCAAAATTTATTTGAAATCCATGTAACACGAAAGCACTTGGGTGAA 63682
QY 1561 ACAAAAGCTGAAGCACTTTTCAACAACGGCTGATGATTTGATATACAGGG 1620
Db 63683 ACAAAAGCTGAAGCACTTTTCAACAACGGCTGATGATTTGATATACAGGG 63742
QY 1621 CTTAAATACGTAACCTAATTTATTAAGAAAGTGAAGCTCTGATGATTTATGCTTTGTA 1680
Db 63743 CTTAAATACGTAACCTAATTTATTAAGAAAGTGAAGCTCTGATGATTTATGCTTTGTA 63802
QY 1681 ACAATTCGATGATTTTGTGTCAATGAAACATATGAGCAACAGAAACCTAAGT 1740
Db 63803 ACAATTCGATGATTTTGTGTCAATGAAACATATGAGCAACAGAAACCTAAGT 63862
QY 1741 TCACCAAGAAAGTACCAAAACAACAGAAAGTACCTTACACAGGTGGCAAGGA 1800
Db 63863 TCACCAAGAAAGTACCAAAACAACAGAAAGTACCTTACACAGGTGGCAAGGA 63922
QY 1801 ATCTACGTTTACTTGAAGAGGGCGAGCTTGTCTAATTTATGCAAGAGCTTACTTTGCT 1860
Db 63923 ATCTACGTTTACTTGAAGAGGGCGAGCTTGTCTAATTTATGCAAGAGCTTACTTTGCT 63982
QY 1861 AGACCTAGAAAAGAAATGCT 1881
Db 63983 AGACCTAGAAAAGAAATGCT 64003

```

```

RESULT 8
LOCUS BD191808 1687 bp DNA linear PAT 17-JUL-2003
DEFINITION Enterococcus Faecalis polynucleotides and polypeptides.
ACCESSION BD191808
VERSION BD191808.1 GI:33001547
KEYWORDS JP 2002516571-A/110.

```

```

SOURCE
  ORGANISM
    synthetic construct
  other sequences; artificial sequences.
REFERENCE
  1 (bases 1 to 1687)
  Kunesh, C.A., Choi, G.H., Bailey, C. and Hromockyj, A.
  Enterococcus Faecalis polynucleotides and polypeptides
  Patent: JP 2002516571-A 110 04-JUN-2002;
  HUMAN GENOME SCIENCES INC
COMMENT
  PN JP 2002516571-A/110
  PD 04-JUN-2002
  PF 04-MAY-1998 JP 1998548289
  PR 06-MAY-1997 US 60/044031, 16-MAY-1997 US 60/046655 PR
  14-NOV-1997 US 60/066009
  PI CHARLES A KUNESH, GIL H CHOI, CAMELIA BAILEY, ALEX HROMOCKYJ PC
  C12N15/31, C07K14/315, C07K16/12, C12Q1/68, C12N1/21, C12N5/12, PC
  G01N33/569
  PC G01N33/68, A61K39/09
  CC Scrandedness: Double;
  CC Topology: Linear;
  FH Key Location/Qualifiers.
FEATURES
  source
    1..1687
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
ORIGIN
  Query Match 89.7%; Score 1687; DB 2; Length 1687;
  Best Local Similarity 100.0%; Pred. No. 0;
  Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  96 AGAAGAAATGGGGAGAGGCGACAGCTGTGTTCACAAAGAAATGACGATTTAC 155
  1 AGAAGAAATGGGGAGAGGCGACAGCTGTGTTCACAAAGAAATGACGATTTAC 60
  156 AGATCCGCTTATTTCAAAATATGCGGAAAGAAATGAGCGATTTGATTAATCAAGACT 215
  61 AGATCCGCTTATTTCAAAATATGCGGAAAGAAATGAGCGATTTGATTAATCAAGACT 120
  216 GGCAGATGTGATTTAGTATTTATTAACGTGACGAACGAATTTTACGACCAACGAGCGGC 275
  121 GGCAGATGTGATTTAGTATTTATTAACGTGACGAACGAATTTTACGACCAACGAGCGGC 180
  276 AGGCGCAAGCTTGTATGACGCTTAACCAAGCTGTCCAAAGTTTAACTCCCTGGGAACTGT 335
  181 AGGCGCAAGCTTGTATGACGCTTAACCAAGCTGTCCAAAGTTTAACTCCCTGGGAACTGT 240
  336 TGCTCAAGAAACCAACGATGCAATGCGAATGTCACTGTTCACTTAACCTAATAAACAATA 395
  241 TGCTCAAGAAACCAACGATGCAATGCGAATGTCACTGTTCACTTAACCTAATAAACAATA 300
  396 TGTAAAGATGACGTGTATACCATTAAGAAAGAACCAAAAGAGGTGTATGCTGTCTAC 455
  301 TGTAAAGATGACGTGTATACCATTAAGAAAGAACCAAAAGAGGTGTATGCTGTCTAC 360
  456 GAAATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 515
  361 GAAATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
  516 ATATGAAACAGAAATATAGCGTTGTTCAATTTATCTTAATAATGTGTGTGTGTGTGTGTGT 575
  421 ATATGAAACAGAAATATAGCGTTGTTCAATTTATCTTAATAATGTGTGTGTGTGTGTGTGT 480
  576 TGTATGTTTACATGTGAAGAAAGTGAAGATGCTGTAAGAAATGAGAAATTAATGCGCGACA 635
  481 TGTATGTTTACATGTGAAGAAAGTGAAGATGCTGTAAGAAATGAGAAATTAATGCGCGACA 540
  636 ATTGTGTTTCTAATAAGCAAGGCTCACCGGCAAGTAAATTAATCAAGAGAGTCA 695
  541 ATTGTGTTTCTAATAAGCAAGGCTCACCGGCAAGTAAATTAATCAAGAGAGTCA 600
  696 AGATGATTAATATCATGACCAACGATTAAGAAACCAACCAACCAACCAACCAACCAACCAAC 755

```


DB 601 AGATGATATATATACATGACAAACGAGTAAAGAACAGCAAAACGTTTATTACTGGAA 660
QY 756 AAGTATATGAAATTCGCGAAATGATTTTCACAGACGAGATTCGAAACGGGAAATTAC 815
DB 661 AAGTTATGAAATTCGCGAAATGATTTTCACAGACGAGATTCGAAACGGGAAATTAC 720
QY 816 AGTTAAAAATCTTGAGGTTGGTTCGATATTTAGAGAGAGTAAAGCTCCAAATATAC 875
DB 721 AGTTAAAAATCTTGAGGTTGGTTCGATATTTAGAGAGAGTAAAGCTCCAAATATAC 780
QY 876 AGAATTAATTTGAAATCAAAACAAACACCATTTTCAATTTGAGCAACCAATCAACACC 935
DB 761 AGAATTAATTTGAAATCAAAACAAACACCATTTTCAATTTGAGCAACCAATCAACACC 840
QY 936 TGTGAAAAAATGATCAAAATGATTCCTTAAAGTGTATTAACACCAAGCTTGA 995
DB 841 TGTGAAAAAATGATCAAAATGATTCCTTAAAGTGTATTAACACCAAGCTTGA 900
QY 996 TGTGAAAAATGAGCAATGGCAATGGCAAAATTTAAATTTCTGTAATATTCCATT 1055
DB 901 TGTGAAAAATGAGCAATGGCAATGGCAAAATTTAAATTTCTGTAATATTCCATT 960
QY 1056 GGGGATTCGACAAAGAAAGCGCTAATAATACGTCAATTTCAATTTAGTTGATA 1115
DB 961 GGGGATTCGACAAAGAAAGCGCTAATAATACGTCAATTTCAATTTAGTTGATA 1020
QY 1116 ACATGATGACGCTTAACTTTGATTAACGTACTTCTGAGAGTATGCTTAAGCTTAA 1175
DB 1021 ACATGATGACGCTTAACTTTGATTAACGTACTTCTGAGAGTATGCTTAAGCTTAA 1080
QY 1176 TGAATGGGATACAGTATGCTCTGAAATTTATCAAGTGAAGTGAACAGCAATGCTT 1235
DB 1081 TGAATGGGATACAGTATGCTCTGAAATTTATCAAGTGAAGTGAACAGCAATGCTT 1140
QY 1236 CACTGTCGCGTTAATCCAGCGTATATTCCTACGCTAACACAGCGGCGACCTAAAT 1295
DB 1141 CACTGTCGCGTTAATCCAGCGTATATTCCTACGCTAACACAGCGGCGACCTAAAT 1200
QY 1296 CGTTTACTTATGATTTTAAAGTGAAGAGCAGATCCTACGAAAGCTTTAAATGAGGC 1355
DB 1201 CGTTTACTTATGATTTTAAAGTGAAGAGCAGATCCTACGAAAGCTTTAAATGAGGC 1260
QY 1356 GAATGTTGATACCGTACATACGACGACCAACCAACCACTGTTGAAGTGTGACAG 1415
DB 1261 GAATGTTGATACCGTACATACGACGACCAACCAACCACTGTTGAAGTGTGACAG 1320
QY 1416 TGGGAAACGTTTCATTTAAAGTGAATGAGTGAAGGACCAAGCTTGGCGGAGC 1475
DB 1321 TGGGAAACGTTTCATTTAAAGTGAATGAGTGAAGGACCAAGCTTGGCGGAGC 1380
QY 1476 TTTCTTTGTCCTCGGTGATCAAAACAGCGACACAGCAATTTATTTGAAATGAGTAA 1535
DB 1381 TTTCTTTGTCCTCGGTGATCAAAACAGCGACACAGCAATTTATTTGAAATGAGTAA 1440
QY 1536 AACGAAAGCAGCACTTGGGTGAAAAACAAACCTGAGCACTACTTTTACCAACGCG 1595
DB 1441 AACGAAAGCAGCACTTGGGTGAAAAACAAACCTGAGCACTACTTTTACCAACGCG 1500
QY 1596 TGAATGATTTAGTTGATACAGGCGTTAAATACGCTTATTTTGAAGAACTGT 1655
DB 1501 TGAATGATTTAGTTGATACAGGCGTTAAATACGCTTATTTTGAAGAACTGT 1560
QY 1656 AGCTCTGATGATTTATGTTTAAACAAATGAGATTTGAGTGAATTTGAGTGAATTT 1715
DB 1561 AGCTCTGATGATTTATGTTTAAACAAATGAGATTTGAGTGAATTTGAGTGAATTT 1620
QY 1716 ATATGCAACAAGAAACCTAGTTTCAACAGAAATGATCAACAAACAAAGGATAC 1775
DB 1621 ATATGCAACAAGAAACCTAGTTTCAACAGAAATGATCAACAAACAAAGGATAC 1680
QY 1776 CTAACT 1782
DB 1681 CTAACT 1687

RESULT 9
AA228553 1687 bp DNA linear PAT 20-DEC-2002
LOCUS
DEFINITION
Sequence 219 from patent US 6448043.
AR228553
ACCESSION
AR228553.1 GI:27267390
VERSION
AR228553.1
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 1687)
AUTHORS
Choi, G.-H., Bailey, C., Hromockyj, A., and Kunsch, C.A.
TITLE
Enterococcus faecalis EF040 and uses thereof
JOURNAL
Patent: US 6448043-A 219 10-SEP-2002;
Human Genome Sciences, Inc.; Rockville, MD;
WOX;
FEATURES
location/Qualifiers
source 1..1687
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 89.7%; Score 1687; DB 2; Length 1687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 96 AGAAGAAATGCGGAGAGCGCACAGCTGCTGATTCACAAAAGAAATGACGGATTTACC 155
DB 1 AGAAGAAATGCGGAGAGCGCACAGCTGCTGATTCACAAAAGAAATGACGGATTTACC 60
QY 156 AGATCGCTTATTTCAATTAAGCGGAGAAATGAGGAGTTTGAATTAATCAAGACT 215
DB 61 AGATCGCTTATTTCAATTAAGCGGAGAAATGAGGAGTTTGAATTAATCAAGACT 120
QY 216 GGCAGATGACGTTAGTATTATTAACGTGACGAACGAATTTTACGACACAGCGGC 275
DB 121 GGCAGATGACGTTAGTATTATTAACGTGACGAACGAATTTTACGACACAGCGGC 180
QY 276 AGGCGCAAGCGTTGATGACGTTAAACAGCTGTCCAAAGTTTAACTCTGGGAACTGT 335
DB 181 AGGCGCAAGCGTTGATGACGTTAAACAGCTGTCCAAAGTTTAACTCTGGGAACTGT 240
QY 336 TGTCTAAGGAACCAACGATGCAATGGAATGGAATGTCTGCTGATTAACCTAAATAA 395
DB 241 TGTCTAAGGAACCAACGATGCAATGGAATGGAATGTCTGCTGATTAACCTAAATAA 300
QY 396 TGTAAAGATGACGATGATACCATTAAGAAAGAACCAAAAGAGGTGATGCTGTAC 455
DB 301 TGTAAAGATGACGATGATACCATTAAGAAAGAACCAAAAGAGGTGATGCTGTAC 360
QY 456 GAATATGATGATGCGCTTCCAGTTTACGAAATGATCAAGCAACAGATGTTCTTADA 515
DB 361 GAATATGATGATGCGCTTCCAGTTTACGAAATGATCAAGCAACAGATGTTCTTADA 420
QY 516 ATATGGAACAGAGAAATTAGGCGTTGTCATTTATCTTAAATATGATGATGCAATGA 575
DB 421 ATATGGAACAGAGAAATTAGGCGTTGTCATTTATCTTAAATATGATGATGCAATGA 480
QY 576 TGTAGATTACATGTGAAGAAATGAGAACTGCTGAAATGAAAGATTAATGCGCGAGA 635
DB 481 TGTAGATTACATGTGAAGAAATGAGAACTGCTGAAATGAAAGATTAATGCGCGAGA 540
QY 636 ATTGTTATTTTAAAGGAAAGCTCACAGGCAAGTAAATATTCAGAGAGTCA 695
DB 541 ATTGTTATTTTAAAGGAAAGCTCACAGGCAAGTAAATATTCAGAGAGTCA 600
QY 696 AGATGATTAATTAATGAGCAAGGATTAAGAAACAAAGCAAAACCTTTATTAAGTGGGA 755
DB 601 AGATGATTAATTAATGAGCAAGGATTAAGAAACAAAGCAAAACCTTTATTAAGTGGGA 660
QY 756 AAGTTATGAAATTCGCGAAATGATTTTCAACAGAACAGAGATGAAACGGGAAATTAA 815


```
Db 661 AAGTTATGAATTCGCGAATAATGATTCACAGAAAGCAGAAATGGAACGGGAGAAATTAAC 720
Qy 816 AGTTAAAAATCTGAGGTTGGTTCGTATATTTTGAAGAGTAAGTAAAGCTCAATATATGC 875
Db 721 AGTTAAAAATCTGAGGTTGGTTCGTATATTTTGAAGAGTAAGTAAAGCTCAATATATGC 780
Qy 876 AGAATTATTAATGAATAATCAAAACAAAACCAATTTAACAATTGAAGCAAAATCAACACC 935
Db 781 AGAATTATTAATGAATAATCAAAACAAAACCAATTTAACAATTGAAGCAAAATCAACACC 840
Qy 936 TGTGTAAGAAAAACGTCAAAATATGATACCTTAAGTTGATTAACAACACCAAGCTTAGA 995
Db 841 TGTGTAAGAAAAACGTCAAAATATGATACCTTAAGTTGATTAACAACACCAAGCTTAGA 900
Qy 996 TGTGTAAGAAAAATGTCGCAATTCGCGAATAATTAATTAATTTCTGTAAATTTCCATT 1055
Db 901 TGTGTAAGAAAAATGTCGCAATTCGCGAATAATTAATTAATTTCTGTAAATTTCCATT 960
Qy 1056 GGGGATTCGACAGACAAAGAGGCGACGTAAATTAATAGTCAAAATTCATTTAGTTGATA 1115
Db 961 GGGGATTCGACAGACAAAGAGGCGACGTAAATTAATAGTCAAAATTCATTTAGTTGATA 1020
Qy 1116 ACATGATGACGCTTAACCTTTGATTAACGTACTTCGAGAGTATGCTTATGCGTTATA 1175
Db 1021 ACATGATGACGCTTAACCTTTGATTAACGTACTTCGAGAGTATGCTTATGCGTTATA 1080
Qy 1176 TGAATGGGATGACGATATTCCTCTGTAATTAATCAAGTATCTGAACACCAATATGCTT 1235
Db 1081 TGAATGGGATGACGATATTCCTCTGTAATTAATCAAGTATCTGAACACCAATATGCTT 1140
Qy 1236 CACTGTCGCGCTTAATCCAGCGTATATTCCTAGCTTAACACGAGGCGCACTAAATTT 1295
Db 1141 CACTGTCGCGCTTAATCCAGCGTATATTCCTAGCTTAACACGAGGCGCACTAAATTT 1200
Qy 1296 CGTTTACTTTATGATTAATGAATAAGAGAGATCCTAAGAAAGCTTTAAATATGAGGC 1355
Db 1201 CGTTTACTTTATGATTAATGAATAAGAGAGATCCTAAGAAAGCTTTAAATATGAGGC 1260
Qy 1356 GAATGTTGATTAACGTCATACCGACGACCAACCAACCACTGTTGAAGTTGTGACAG 1415
Db 1261 GAATGTTGATTAACGTCATACCGACGACCAACCAACCACTGTTGAAGTTGTGACAG 1320
Qy 1416 TGGGAAAACGTTTCATTAATGAATGCGATGCGATGTAAGGACCAAGCCCTGGCGGAGC 1475
Db 1321 TGGGAAAACGTTTCATTAATGAATGCGATGCGATGTAAGGACCAAGCCCTGGCGGAGC 1380
Qy 1476 TTCTTTTGTGTCGTCGTCATCAAAACAGCGACACAGCAATTAATTTGAATGATGAAC 1535
Db 1381 TTCTTTTGTGTCGTCGTCATCAAAACAGCGACACAGCAATTAATTTGAATGATGAAC 1440
Qy 1536 AACGAAAGCAGCACTTGGGTGTAACAAAGCTGAAGCACTACTTTTACACCAAGCGC 1595
Db 1441 AACGAAAGCAGCACTTGGGTGTAACAAAGCTGAAGCACTACTTTTACACCAAGCGC 1500
Qy 1596 TGAATGATTAATGATTAATCAAGGCTTAATTAACGCTAATTTTGAAGAACTGT 1655
Db 1501 TGAATGATTAATGATTAATCAAGGCTTAATTAACGCTAATTTTGAAGAACTGT 1560
Qy 1656 AGCTCTGATGATTAATGCTTGTGTAACAATCGATGTAATTTGTCATGAAACAATC 1715
Db 1561 AGCTCTGATGATTAATGCTTGTGTAACAATCGATGTAATTTGTCATGAAACAATC 1620
Qy 1716 AATATGACACACGAAACCTTGTTTACACGAAACCAACCAACCAACCAAGCTTAC 1775
Db 1621 AATATGACACACGAAACCTTGTTTACACGAAACCAACCAACCAACCAAGCTTAC 1680
Qy 1776 CTTACCT 1782
Db 1681 CTTACCT 1687
```

RESULT 10

```
AR692968
LOCUS AR692968 1687 bp DNA linear PAT 14-SEP-2005
DEFINITION Sequence 219 from patent US 6913907.
ACCESSION AR692968
VERSION AR692968.1 GI:75182711
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1687)
AUTHORS Chol, G.H., Bailey, C., Hromockyj, A. and Kunsch, C.A.
TITLE Enterococcus faecalis polynucleotides encoding EF059
JOURNAL Patent: US 6913907-A 219 05-JUN-2005;
Human Genome Sciences, Inc.; Rockville, MD;
MOX;
FEATURES
source location/Qualifiers
1..1687
/orfname="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 89.7%; Score 1687; DB 2; Length 1687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 96 AGAAGAAATGGGAGAGGCGACAGCTCGTGAATTCACAAAAGAAATGACGATTTAC 155
Db 1 AGAAGAAATGGGAGAGGCGACAGCTCGTGAATTCACAAAAGAAATGACGATTTAC 60
Qy 156 AGATCCGCTTATTAATAGCGGGAAGAAATGAGCGAGTTGATTAATTAATCAAGACT 215
Db 61 AGATCCGCTTATTAATAGCGGGAAGAAATGAGCGAGTTGATTAATTAATCAAGACT 120
Qy 216 GCGAGATGACGCTTATGATTTATTAACGTGACGAACGAATTTTACGACCAACGAGCGC 275
Db 121 GCGAGATGACGCTTATGATTTATTAACGTGACGAACGAATTTTACGACCAACGAGCGC 180
Qy 276 AGGCGCAAGCGTTGATGACGCTTAACCAAGTCTCCAAAGTTTATCTCCGGAACCTGT 335
Db 181 AGGCGCAAGCGTTGATGACGCTTAACCAAGTCTCCAAAGTTTATCTCCGGAACCTGT 240
Qy 336 TGCCTAAGGAACCAACGATGCAATGGGAATGTCATCTGTCAGTTACTTAATAACAAA 395
Db 241 TGCCTAAGGAACCAACGATGCAATGGGAATGTCATCTGTCAGTTACTTAATAACAAA 300
Qy 396 TGGTAAAGATGACGTATACCATTAAGAAACCAAAAGAGGTGTATGCTGTAC 455
Db 301 TGGTAAAGATGACGTATACCATTAAGAAACCAAAAGAGGTGTATGCTGTAC 360
Qy 456 GAATATGCTGTGTCGCTTCCAGTTTGAATGATCAAGCAACAGATGTTCTATTA 515
Db 361 GAATATGCTGTGTCGCTTCCAGTTTGAATGATCAAGCAACAGATGTTCTATTA 420
Qy 516 AATATGAAACAGAAATAGCGGTGTCATATTAATCCTTAATAATGTGTGTCGAATGA 575
Db 421 AATATGAAACAGAAATAGCGGTGTCATATTAATCCTTAATAATGTGTGTCGAATGA 480
Qy 576 TGTATGTTTACATGTGAAAAAAGTAGAACTGCTGAAATGAAAGATTAATGCGCAGA 635
Db 481 TGTATGTTTACATGTGAAAAAAGTAGAACTGCTGAAATGAAAGATTAATGCGCAGA 540
Qy 636 ATTGTATTTTCTTAATACGGAAGGCTCACAGGCAAGTAAATATATTCAGAGAGTCA 695
Db 541 ATTGTATTTTCTTAATACGGAAGGCTCACAGGCAAGTAAATATATTCAGAGAGTCA 600
Qy 696 AGATGATTAATATACATGACCAACGATTAAGAAACCAACCAACCAACCAAGCTTATTA 755
Db 601 AGATGATTAATATACATGACCAACGATTAAGAAACCAACCAACCAACCAAGCTTATTA 660
Qy 756 AAGTTATGAATTTGGCGAAATGATTTCAAGAGCAGAGATGACGGAGAAATTAAC 815
Db 661 AAGTTATGAATTTGGCGAAATGATTTCAAGAGCAGAGATGACGGAGAAATTAAC 720
```

```
QY 816 AGTAAAAATCTTGAGGTGTTGCTATATTTTAAAGAAAGTAAAGCTCAAAATATGCG 875
DB 721 AGTAAAAATCTTGAGGTGTTGCTATATTTTAAAGAAAGTAAAGCTCAAAATATGCG 880
QY 876 AGAATTAATGAAAAATCAAAACAAAAACCAATTTTCAATTTGAAGCAACATCAACACC 935
DB 781 AGAATTAATGAAAAATCAAAACAAAAACCAATTTTGAAGCAACATCAACACC 840
QY 936 TGTGAAAAAACAAGTCAAAAAATGATTCCTTAAAGTTGATTAACCAACCAAGCTTGA 995
DB 841 TGTGAAAAAACAAGTCAAAAAATGATTCCTTAAAGTTGATTAACCAACCAAGCTTGA 900
QY 996 TGTGAAAAATGTCGCAATTTGGCAAAAAATTAATCAATTTCTGTAATATTTCCATT 1055
DB 901 TGTGAAAAATGTCGCAATTTGGCAAAAAATTAATCAATTTCTGTAATATTTCCATT 960
QY 1056 GGGGATTTGCGACAAAGAAAGCGCGCTAATAAATACGTCAAATTTCAATTTAGTTGATA 1115
DB 961 GGGGATTTGCGACAAAGAAAGCGCGCTAATAAATACGTCAAATTTCAATTTAGTTGATA 1020
QY 1116 ACATGATGACGCTTAACTTTGATGATACGTGACTTTGAGAGATGCTTATGCGTTATA 1175
DB 1021 ACATGATGACGCTTAACTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1176 TGTGAGGAGTACAGTGAATGCTCTGAAATTTATCAAGTGAAGTGAAGCAAGATGCTT 1235
DB 1081 TGTGAGGAGTACAGTGAATGCTCTGAAATTTATCAAGTGAAGTGAAGCAAGATGCTT 1140
QY 1236 CACTGTCGCGTTAATCCAGCGTATATTCCTAAGCTTAAACCAAGCGCGACACTAAAT 1295
DB 1141 CACTGTCGCGTTAATCCAGCGTATATTCCTAAGCTTAAACCAAGCGCGACACTAAAT 1200
QY 1296 CGTTTAACTTAACTTAAATGAAAGCAAGATCCTAAGAAAGCTTAAATGAAAGCG 1355
DB 1201 CGTTTAACTTAACTTAAATGAAAGCAAGATCCTAAGAAAGCTTAAATGAAAGCG 1260
QY 1356 GAATGTTGATTAACGCTCATACGACGACCAACCAACCACTGTTGAAGTTGACAG 1415
DB 1261 GAATGTTGATTAACGCTCATACGACGACCAACCAACCACTGTTGAAGTTGACAG 1320
QY 1416 TGGGAAACGTTTCAATTAAGTCGATGGCGATGTCAGGACCAACAGCTTGGCGGAGC 1475
DB 1321 TGGGAAACGTTTCAATTAAGTCGATGGCGATGTCAGGACCAACAGCTTGGCGGAGC 1380
QY 1476 TTTCTTTGTCGTCGCGATCAAAACGACGACCAAGAAATTTGAAATGATGTAAC 1535
DB 1381 TTTCTTTGTCGTCGCGATCAAAACGACGACCAAGAAATTTGAAATGATGTAAC 1440
QY 1536 AACGAAGCAGCACTTGGGTGAAAAACAAAGCTGAAGCACTACTTTTACCAACAGCG 1595
DB 1441 AACGAAGCAGCACTTGGGTGAAAAACAAAGCTGAAGCACTACTTTTACCAACAGCG 1500
QY 1596 TGAATGATTAAGTTGATATACAGGCGTTAATACGCTTATTTTGAAGAACTGT 1655
DB 1501 TGAATGATTAAGTTGATATACAGGCGTTAATACGCTTATTTTGAAGAACTGT 1560
QY 1656 AACTCCTGATGATTAAGTTGTTTAAACAAATGGAATTTGTTGTTGTTGTTGTTGTTGTT 1715
DB 1561 AACTCCTGATGATTAAGTTGTTTAAACAAATGGAATTTGTTGTTGTTGTTGTTGTTGTT 1620
QY 1716 ATATGCGACAAAGAAACCTAGTTTACACAGAAAAAGTACCAACAAACAGAAAGTAC 1775
DB 1621 ATATGCGACAAAGAAACCTAGTTTACACAGAAAAAGTACCAACAAACAGAAAGTAC 1680
QY 1776 CTTTACT 1782
DB 1681 CTTTACT 1687
```

```
RESULT 11
CQ969772 1878 bp DNA linear PAT 20-DEC-2004
LOCUS CQ969772
DEFINITION Sequence 347 from Patent WO2004106367.
```

```
ACCESSION CQ969772
VERSION CQ969772.1 GI:56743733
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLES
JOURNAL
FEATURES
source
location/Qualifiers
/organism="Enterococcus faecium"
/mol_type="unassigned DNA"
/db_xref="taxon:1352"
ORIGIN
Query Match 53.9%; Score 1014.6; DB 2; Length 1878;
Best Local Similarity 71.9%; Pred. No. 21e-290;
Matches 1342; Conservative 0; Mismatches 519; Indels 6; Gaps 1;
QY 1 ATGAAGCAATTAATAAAGTTGTGACCCGTTAGTACCTTGTACTAATTTTGCACCT 60
DB 1 ATGAAGCAATTAATAAAGTTGTGACCCGTTAGTACCTTGTACTAATTTTGCACCTA 60
QY 61 TTCAAGTGTGTTAGGACCAACATGCTCATTTGCGAAGAAATGGGAGAGCGCACAG 120
DB 61 TTCAAGTGTGTTAGGACCAACATGCTCATTTGCGAAGAAATGGGAGAGCGCACAG 120
QY 61 CTCACAAACAGCTTGGGCGCAAAAAGTGTGACAGGAGACAG-----CAGCTCAA 114
DB 61 CTCACAAACAGCTTGGGCGCAAAAAGTGTGACAGGAGACAG-----CAGCTCAA 114
QY 121 CTGTGATTCACAAAGAAAGAAATGACGGAATTTACAGATTCGCTTATTCAAATAGCGGG 180
DB 115 GTGATCTTTCATTAAGAAAGAAATGACGGAATTTACCGGATCTTATTCAAAGACGGG 174
QY 181 AAAAGAAATGAGCGATTTGATTAATATCAAGGACTGGCAATGATGACGTTAGTAT 240
DB 175 AAAAGAAATGAGCGATTTGATTAATATCAAGGACTGGCAATGATGATGATGATGAT 234
QY 241 AACGTGACGAAGATTTTACAGACAGAGGAGCGCAAGGTTGATGCACTAA 300
DB 235 AACGTGACGAAGATTTTACAGACAGAGGAGCGCGTGGGATGACACAA 294
QY 301 CAAGCTGTCGCAAGTTTAACTCTGGGAAACCTGTTGCTCAAGAGACCAAGATCAAT 360
DB 295 CAAGCTGTCGCAAGTTTAACTCTGGGAAACCTGTTGCTCAAGAGACCAAGATCAAT 354
QY 361 GGGAAATGTCAGTTCAGTTACTTAATAAACAATGTTAAAGTGAAGTGTATCCATT 420
DB 355 GGGAAATGTCAGTTCAGTTACTTAATAAACAATGTTAAAGTGAAGTGTATCCATT 414
QY 421 AAAAGAAACCAAAAGAGGAGTGTAGTGTGCTAGAAATGATGAGGCGTTCCAGTT 480
DB 415 AAAAGAAACCAAAAGAGGAGTGTAGTGTGCTAGAAATGATGAGGCGTTCCAGTT 474
QY 481 TACGAATGATCAACCAACAGATGTTCTTAATATGAAACAGAAATTAAGCGTT 540
DB 475 TATGATGATCAACCAACAGATGTTCTTAATATGAAACAGAAATTAAGCTACT 534
QY 541 GTTCATATTTATCTTAATAATGTGTAGCAATGATGTTAGTTTACATGTGAAAAAGTA 600
DB 535 ATCCATCTCTACCTTAATAATCACTGCTAATGATGAAACCTTAAGTTTCAAAAAATC 594
QY 601 GGAATGTCGAAAGAAAGGATTAATGAGCGGAGAAATTTGATTTCTTAAGAGAGCG 660
DB 595 GGAATGTCGAAAGAAAGGATTAATGAGCGGAGAAATTTGATTTCTTAAGAGAGAG 654
QY 661 TCACGAGGACAGTAAATATATCCAGGAGTCAAGATGATGATTAATATCATGACAG 720
DB 655 ACACGAGGTCGAAAGAAATATCCAGGAGTCAAGATGATGATGATGATGATGATGATGAT 714
QY 721 GATTAAGAACAGCAACAGCTTTTATGATGAGGAAAGTATGAAATTTGGCAAAATGAT 780
DB 714 GATTAAGAACAGCAACAGCTTTTATGATGAGGAAAGTATGAAATTTGGCAAAATGAT 780
```

```

Db      715 GATCAACCAAGCCAAACATTTCATTACCTGCTGATTCCTTATGACATCGGCAACATGAC 774
Qy      781 TTCAAGAGAGCAGAGATGGAACGGAGAAATTACGTTAAAAATTTAGAGTTGGTTGG 840
Db      775 TTTGCCGAGGCACTCATGAAAAAGGCCAGTTGATGATTATCATTTAGAAAGTTGGAAAA 834
Qy      841 TATATTTTGAAGAGTAAAGCTCCAAATTAATGAGAAATTAATTGAAATCAACAAAA 900
Db      835 TATTAATTTGAAAGAGTAAAGCTCTGATTAATGCGAAATGATTTGAAAGCAACATC 894
Qy      901 ACACCATTTTACATTTGAGCAAAACATCAACACCTGTTGAAAAAAGCACTCAAAAATGAT 960
Db      895 AGCCCTTTTGAAGATCCTGCAATAGCCAAACCCAGTGAAGAAAAGCATTCAAAAATGAT 954
Qy      961 ACCCTTAAAGTTGATTAACCAACACCAAGCTTAGATGATGAAGATTTGGCAATTGGCGAA 1020
Db      955 AGCTCAAAAGTTGATTAACCAACACCTCAATTTGAAGAAAAGATTCCTCAATCCGCTGAA 1014
Qy      1021 AAAATTTAATATCAATTTCTGTAATATTTCCATTTGGGAGTTGACAGCAAAAGGCGAC 1080
Db      1015 AAAATTTCAATATGAGATTTCTGCAATATCCCATTTAGATGATGATTAAGAGAGAACG 1074
Qy      1081 GCTAATTAATACGTCATTAATTCATTTAGTTGATTAACATGATGACACCTTTAACTTTGAT 1140
Db      1075 CAACCAAGTACACCAACATTTCAAACTTATGATGATCATGACGCTGCTTTAACATTTGAT 1134
Qy      1141 AACGTGACTTCTGAGAGATGATGCTTATGCTTATATGATGGGATTCAGTGAATGCTCCT 1200
Db      1135 AATGATTTCTTCAAGAACTGATCTTATGCTTATATGATGAATTAAGAAATTCGACCCA 1194
Qy      1201 GAAATTTATCAAGTGTGTAAGCAAGCAAAAGGCTTCACTGTCGCGCTTATCAAGGCTAT 1260
Db      1195 GTAAATTTATTTCTGTCTGATGAAACAGACGATTCACGTTTCACTGATTCGAAATTAAT 1254
Qy      1261 ATTCTACGCTTAAACCAAGCGGCACTAATAATTCGTTACTTTATGCAATTTAATGA 1320
Db      1255 ATTCCTTATTAATCTCTGGCGGTCATTGAATTCGTTTATCATATGCAATTTGAACGAA 1314
Qy      1321 AAAGCATCTTACGAAAGGCTTTAAAAATGAGCGAATGTTGATTAAGCTTATCCGAC 1380
Db      1315 AAAGCATCTTACGAAAGGCTTTCTAAACCAAGCAAAAGTGCATTAACGCGCATCAAAAT 1374
Qy      1381 GACCAAAACCAACCACTGTTGAAGTTGACAGGCGGAAAGCTTCAATTAAGTCAAT 1440
Db      1375 GATCAAAACCAACCGTCACTGATGATGCTGTTACTGGGCGCAACGATTTGTTAAAGTAT 1434
Qy      1441 GCGCATGTCACAGCAGCAACAGCCTTGGCGGAGCTTCTTGTGCTCGCTGATCAAAAC 1500
Db      1435 GGTGACGTTACATCAGACCAACATCTGTCGAGAGCAAGATTCGTCGTGATCAAGAT 1494
Qy      1501 AGCGACACAGCAATTTATTTGAAATTCATGAAACAGCAAAAGCAGCACTTGGGTGAAA 1560
Db      1495 AGTGACACAGCAATTTATTTATCGATCGACCCATCCCAAAAGCCCTCAGCTGGGATCG 1554
Qy      1561 ACAAAGCGTGAAGCACTACTTTTCAACAAGCGGCTGATGATTTGTTATATCAAGG 1620
Db      1555 GCGAAAGATCAGCAACGCTTTTACACACCAAGTAAACGTTTAACTGATGACAGAT 1614
Qy      1621 CTTAATAACGTAACCTATTTATTAAGAAACCTGAGCTCTGATGATTTATGCTTTGTTA 1680
Db      1615 CTAAATATTTGGACGATCTATCTGAAAGAAAGAAAGCGCCAGAAATATGTTCCATTA 1674
Qy      1681 ACAATTCGATTTGATTTGTGCTCAATGAAACATCATATGCGACACAGAAACCTAGTT 1740
Db      1675 ACAAAACCGTGTGATTTCTATCTGATGAAACATCTTATGTAACAGCAAGACGTTGAT 1734
Qy      1741 TCACCAAGAAAGTACCAACAAACCAAGGACCTTACCTTCAACAGATGGCAAAAGGA 1800
Db      1735 TCTCTGAAAGAAATACCAATTAACCAAGGATCACTTCTTCAACAGCGGTGAAAGGA 1794
Qy      1801 ATCTACGTTTACTTAGAAGTGCGCGAGCTTCTGCTAATTATGACAGAGCTTACTTGTCT 1860
Db      1795 ATCTATGTGTATATCGGTGACAGAGATATCTTCTATCTGATTTGCTGACATGATCTTGTCT 1854

```

```

Qy      1861 AGACGTA 1867
Db      1855 AGACGCA 1861

RESULT 12
AR348987
LOCUS      AR348987      1995 bp      DNA      linear      PAT 17-AUG-2003
DEFINITION Sequence 3598 from patent US 6583275.
ACCESSION AR348987
VERSION    AR348987.1 GI:33746032
KEYWORDS
SOURCE
ORGANISM   Unknown.
            Unclassified.
REFERENCE  1 (bases 1 to 1995)
            Doucette-Stamm, L.A. and Bush, D.
            Nucleic acid sequences and expression system relating to
            Enterococcus faecium for diagnostics and therapeutics
            Patent: US 6583275-A 3598 24-JUN-2003;
            Genome Therapeutics Corporation; Waltham, MA
FEATURES
            source
            1..1995
            /organism="Unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      53.9%; Score 1014.6; DB 2; Length 1995;
Best Local Similarity 71.9%; Pred. No. 2.1e-290;
Matches 1342; Conservative 0; Mismatches 519; Indels 6; Gaps 1;

Qy      1      ATGAAGCAATTAATAAAGTTGTTGTAACCCGTTAGTACTTGTACTAATTTTGCACCTT 60
Db      118      ATGAATAATCATTAATAAATAAATTAACGTTATGTAGAGTCTTTTCTTATTTTACATTA 177
Qy      61      TTCACAAGTGTATTTAGGCAACAACGTCATTTGCAAGAAATATGGGAGAGCGACAG 120
Db      178      CTCACAAACAGCTTGGCGCAAAAAGTTGTCAGAGAGACAG-----CAGCTCAA 231
Qy      121      CTCGTGATTCACAAAAGAAAATGAACGATTTTACAGATTCGCTTATTTCAAAATAGCGG 180
Db      232      GTCATCTTCATTAATAAAGAAAATGACTGATTTACCGATCTTTATCCAAAACAGCGGG 291
Qy      181      AAAAAGATGAGCAAGTTGATTAATTCAGAGATCGGCAAGTGAACGTTTATTAATTT 240
Db      292      AAAAAGATGAGCAATTCATCAATACCAAGATTTAGCCGATATTTTCAATTTTCAAGTTAT 351
Qy      241      AACGTGACGAAGATTTTATCGAGCAACGAGCGGCGCAAGCGTGTATGATGACGCTAAA 300
Db      352      AACGTCACTCAAGAAATTTTATGCGCAAGAGATTAAGAGCGTCCGTGATGACGCAAAA 411
Qy      301      CAAAGTGTCCAAAGTTTAACTCTCGGAAAACCTGTTGCTCAAGAAACACCGATGCAAT 360
Db      412      CAAAGATGTCAGTCTTTGATCTCTGATACACAGATTTGCTTCAAGAAACAGATGCTGAT 471
Qy      361      GGAATATGTCAGTCTTCAAGTAAACAAATATGTAAGATGATGATGATTAATTTACCAT 420
Db      472      GGAATATGTCAGTCTTATCTTAACTTAAACAAATATGGAAGATGCACTCAACGATC 531
Qy      421      AAAAGAAACCAAAAGAGGAGTGTAGTGTGCTACGAATATGATGATGATGATGATGAT 480
Db      532      AAAAGAAACCAAAAGAGGAGTGTAGTGTGCTACGAATATGATGATGATGATGATGAT 591
Qy      481      TAGAATATGATCAAGCAACAGATGTTCTTAAATATGAAACAGAAATTAAGCGGTT 540
Db      592      TATGATATGATCAACAGCAGATGCTTAAATATGAGGACAGAAAGATCAAGTACT 651
Qy      541      GTTCAATTTATCTTAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db      652      ATCCATCTCTACCTTAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 711
Qy      601      GGAATGCTGAATAAGAAAGATTAATATGCGCAGAAATTTGTTATTTCTAAAGCGAAGCG 660

```

```
Db 712 GGTACTGCGGAAAACGAACTAAATGAGAGAAATTTATTTCTTAAAGAAAGGA 771
Qy 661 TCACGAGGACAGTAATAATATCCAAAGAGTCAAAAGATGATATATATCATGAGACAG 720
Db 772 ACACCAAGGCGTAAATAATATCATCCAAAGTGCAGATGATGATATGATACCTTGGAACA 831
Qy 721 GATTAAGAAACAAAGCAAAAGCTTTATTTACTGGGAAAAGTTATGAAATTTGGCAAAATGAT 780
Db 832 GATCAAAACCAAGCCAAACATTTTCATTTACTGGTCAATCTTATGACATCGGCAACATGAC 891
Qy 781 TTCAAGAAAGAGAGATGAAACGAGAGATTTAAACAGTTAAATCTTAGGTTGGTGG 840
Db 892 TTTCGCGAGGCAATCTATTTGAAAAAGGCCAGTTGATGTTATATCATTTAGAAAGTTGAAAA 951
Qy 841 TATATTTTGAAGAGATTAAGCTCCAAATATGCAAAATTAATGAAATCAAAACAAA 900
Db 952 TATATTTTGAAGAGATTAAGCTCCGATTAATGCGGAAATGATGAAAGCAAAACAATC 1011
Qy 901 ACACCTTTACATTTGAAACCAATCAAAACCTGTTGAAAAAAACAGTCAAAATGAT 960
Db 1012 AGGCTTTTGAAGTCTGGCAAAATAGCCAAACACAGTAGAAAAACCAATCAAAATGAT 1071
Qy 961 ACCTTAAAGTTGATTAACCAACACCAAGCTTAGATGTTAAAGATGTGGCAATTTGGCGAA 1020
Db 1072 AGCTTAAAGTTGATTAACCAACACCTCAATTTGAATGAAAAAGATGTGGCAATCGGTGAA 1131
Qy 1021 AAAATTAATTCAAATTTCTGTAATATTCATTTGGGAGATTTGCAGACAAAGAGCGAC 1080
Db 1132 AAAATTAATTCAGATTTCTGTCATATTCCTCATTTAGATATGCTGATTAAGAAAGGAAG 1191
Qy 1081 GCTAATTAATTCAGTCAATTTCAATTTAGTTGATTAACATGATGACCTTAACTTTTGAAT 1140
Db 1192 CAACCAAGATCAACCAATTTCAATTTAGTATGATGACGCTCTTAAACATTTGAT 1251
Qy 1141 AACGTGACTTTGAGAGATATGCTTATGCTTATATGATGAGGAGATACAGTATTTGCTCT 1200
Db 1252 AATGATTTCTTGAGAAAGATATGCTTATGCTTATATGATGAGAAATTAAGAAATCGACCA 1311
Qy 1201 GAAATTAATTCAGTCAAGCAACAGCAAAATGCTTCACTGTGCGCTTATTCACAGCTAT 1260
Db 1312 GTAAATTAATTCGTCTGCTGAGCAAAAGAGATTCACGCTTCACTTATCCGAATTAAT 1371
Qy 1261 ATTCAGAGCTTAACACGAGCGGCAACATTAATGCTTACTTATGATTTAAATGAA 1320
Db 1372 ATTCCTTCACTTAACCTCTGCGGATCATTTGAATTTGTTTACATTAATGATTTGAACGAA 1431
Qy 1321 AAAGCAGATCTTAACGAAAGCTTTAAATAATGAGCGCAATGTTGAACGCTCATCCGAC 1380
Db 1432 AAAGCAGATCTTAACGAAAGCTTTTCTAACAAAGCAATGCTGATTAACGCGCATTCAAAT 1491
Qy 1381 GACCAAAACCAACCAACCTTTGAGTGTGACAGGTGGGAAACGTTTCACTTAAAGTCAT 1440
Db 1492 GATCAAAACCAACCGCTGATGATGCTTAACTGGGGGCAAAAGATTTGTTAAAGTAGAT 1551
Qy 1441 GCGGATGTGACAGCGACCAAGCTTGGGGGAGCTTCCCTTGTGCTCGGTATCAAAAC 1500
Db 1552 GGTGACGTTATCATCAACCAACCTTGTGAGAGCAAAATGCTGCTTGTGATCAAGAT 1611
Qy 1501 AGCGACACAGCAAAATTTATTTGAAATTCATGAAACCAACGAAAGCAACAACTTGGGTGAA 1560
Db 1612 AGTGAACAGCGCAAAATTTATTCATGATCGAACCAATCAAAAGCGCTGAGCTGGGATATG 1671
Qy 1561 ACAAAAGCTGAAGCAACTACTTTTACAAACAGCGCTGATGATTAAGTTATCATACAGG 1620
Db 1672 GCGAAGGATCAAGCAACGCTTTTACAAACCAAGTAGAGTAAATCATGTGACAGAT 1731
Qy 1621 CTAAATTAAGTACATTAATTTAGAAAGAACTGATGCTCGATGATTAATGCTTGTGTA 1680
Db 1732 CTAAATTAAGTACATTAATTTAGAAAGAACTGATGCTCGATGATTAATGCTTGTGTA 1791
Qy 1681 ACAAATCGATTAATTTGTGCTGATGAAACATTCATGATGCAACAGAAAACTTAGTT 1740
```

```
Db 1792 ACAAAACGCTGATGATTTACTATGATGACATCTTATGTAACAGCAGAGAGTTGAT 1851
Qy 1741 TCACGAAAAAGTACCAACCAACCAAGGATCTTACTTCAACAGGTGGCAAGGA 1800
Db 1852 TCTCTGAAAAAATACCAATTAACCAAGGATACATCTTCTTCAACAGCGGTAAAGGA 1911
Qy 1801 ATCTAGTTTACTTGTGGAAGTGGGACAGCTTGTCTATCTTATGAGAGAGTACTTGTCT 1860
Db 1912 ATCTATGTTATATGAGGTGAGAGAGTACTTCTTATGATGATGAGCTGATCTTGTCT 1971
Qy 1861 AGACGTA 1867
Db 1972 AGACGCA 1978

RESULT 13
AX476900 2199 bp DNA linear PAT 12-AUG-2002
LOCUS AX476900
DEFINITION Sequence 28 from Patent W00212294.
ACCESSION AX476900
VERSION AX476900.1 GI:22216159
KEYWORDS
SOURCE Enterococcus faecalis
ORGANISM Enterococcus faecalis
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;
Enterococcus.

REFERENCE
1 Adderson, E. and Bohnsack, J.
AUTHORS Group b streptococcus polypeptides nucleic acids and therapeutic
TITLE compositions and vaccines thereof
JOURNAL Patent: WO 0212294-A 28 14-FEB-2002;
ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US) ; University of Utah
Research Foundation (US)

FEATURES
source location/Qualifiers
1..2199
/organism="Enterococcus faecalis"
/mol_type="unassigned DNA"
/db_xref="taxon:1351"

ORIGIN
Query Match 31.6%; Score 595; DB 2; Length 2199;
Best Local Similarity 100.0%; Pred. No. 1.9e-165;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1287 ACTAAATTCGTTTACTTTATGATTAATGAAAGAGATCCTAGAAAGGCTTTA 1346
Db 1 ACTAAATTCGTTTACTTTATGATTAATGAAAGAGATCCTAGAAAGGCTTTA 60
Qy 1347 AATGAGCGAATGTGATTAACGTCATPACGACGACCAACACCACTGTTGAAGT 1406
Db 61 AATGAGCGAATGTGATTAACGTCATPACGACGACCAACACCACTGTTGAAGT 120
Qy 1407 TGTGACAGGTGGAACGTTTCATTTAAAGTCATGAGCGATGTGACAGGACCAAGCTT 1466
Db 121 TGTGACAGGTGGAACGTTTCATTTAAAGTCATGAGCGATGTGACAGGACCAAGCTT 180
Qy 1467 GCGGAGCTTCCTTTCGTCGCGATCAAAACGCGACACAGCAAAATTAATTTGAAAT 1526
Db 181 GCGGAGCTTCCTTTCGTCGCGATCAAAACGCGACACAGCAAAATTAATTTGAAAT 240
Qy 1527 CGATGAACCAAGCAACCTTGGGTGAAAAACAAAGCTGAAGCAACTACTTTTAC 1586
Db 241 CGATGAACCAAGCAACCTTGGGTGAAAAACAAAGCTGAAGCAACTACTTTTAC 300
Qy 1587 AACACGCTGATGATTTAGTTGATATCAACAGGCTTAAATTCGTTACTTATTTAGA 1646
Db 301 AACACGCTGATGATTTAGTTGATATCAACAGGCTTAAATTCGTTACTTATTTAGA 360
Qy 1647 ACAAACGATGCTCCGATGATTAATGATGCTGTTAAACAAATGGAATTTGGTGA 1706
Db 361 ACAAACGATGCTCCGATGATTAATGATGCTGTTAAACAAATGGAATTTGGTGA 420
Qy 1707 TGAACATCATATGACCAACAGAAACCTAGTTTCACAGAAAAAGTACCAACAAACA 1766
```

Db 421 TGAACATCATATGCGACACAGAAAGAACTAGTTCCAGAAAGAAAGTCCAAACAAATA 480

Qy 1767 CAAGGATACCTTACCTTCAACAGTGGCAAGAAATCTACGTTTACTTGAAGTGGCC 1826

Db 481 CAAAGGATACCTTACCTTCAACAGTGGCAAGAAATCTACGTTTACTTGAAGTGGCC 540

Qy 1827 AGCTTCTGCTACTTATTGTCAGAGTCTACTTGTCTAGAGCTAGAAAGAAATGCT 1861

Db 541 AGCTTCTGCTACTTATTGTCAGAGTCTACTTGTCTAGAGCTAGAAAGAAATGCT 595

RESULT 14
AC181263/c 150038 bp DNA linear HTG 01-FEB-2006

LOCUS Strongylocentrotus purpuratus clone R3-4012A17, WORKING DRAFT

DEFINITION AC181263

ACCESSION AC181263.1 GI:86263499

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_POOLED_CLONE; HTGS_ENRICHED.

SOURCE Strongylocentrotus purpuratus

ORGANISM Strongylocentrotus purpuratus

REFERENCE 1 (bases 1 to 150038)
Shen, Y., Abraham, K.K., Abulimiti, A., Adams, C.O., Aduba, G., Allen, C.C., Alsbrooks, S.L., Anosike, U.S., Archer, P.M., Arrondo, H.H., Attaway, T., Bandaranaike, D.P., Bangura, L., Barton, S.D., Bell, A.V., Bell, S.N., Beraducci, A.R., Bickham, C., Biewald, K., Blyth, P.R., Bunay, C.J., Canada, A., Cardenas, V., Carter, K., Chacko, J., Chandrasekhar, M.N., Chavez, A., Chavez, D., Chen, G., Chen, R., Chu, H., Clerc, blankenburg, K.P., Cockrell, R., Cooper, J., Coyle, M.D., Cree, A., Cueto, C.B., Curry, S.M., Dai, W., Dao, M.D., Davila, M., Davis, C., Davy-Carroll, L., Del fierro, P., Densen, R., Denson, S., Ding, Y., Dinh, H.H., Donlin, J.E., Dugan-Rocha, S., Dunn, A.M., Durbin, K.J., Eboong, V.E., Egan, A., Espinosa, V.C., Fa, M., Fernandez, S., Fernando, P.R., Ferrer, A.R., Flagg, N., Forbes, L.D., Fowler, R.G., Fu, Q., Fuh, E., Gabisi, R.A., Gargadanan, M., Garner, J., Ghose, S., Gargas, M., Gonzalez-Garay, M.L., Guevara, W.V., Haaland, W.C., Heberlein, K.A., Hagans, B.J., Hall, O., Hamid, H., Hamilton, K.A., Hampton, O.A., Haines, B.A., Harris, R.A., Havlak, P., Hawes, A.C., Hawkins, E.S., Haynes, S.J., Hemphill, L., Hernandez, J., Hines, S., Hirani, K., Hitchens, M.E., Hodgson, A.V., Hognes, M.E., Holder, M., Hollins, B., Howell, L.L., Hulik, S.W., Hume, J., Jackson, A., Jackson, L.R., Jacob, S.K., Jhangiani, S.N., Jiang, H., Johnson, B., Johnson, R., Joseph, V., Joy, C., Kalkat, P.B., Kalafatis, K.J., Kalu, J.B., Kang, Y., Keebler, J., Khan, Z.M., Kidwai, S., King, L.M., Kisamo, H., Kovar, F., Kowals, A.N., Kowis, C.R., Lago, L.A., Lago, M.T., Lai, C., Lata, F., Le, T.T., Lee, S.L., Lee, T.W., Legall, J.L., F.H., Lemon, S.J., Lewis, L.R., Li, B., Li, Y., Li, Z., Linnell, M.A., Liu, J., Liu, W., Liu, Y., Liu, Y., Llyanage, D., London, P., Lopez, J., Lorenshewe, L.M., Lozano, R.J., Luc, T., Madu, R.C., Maheshwari, M., Maheshwari, R., Malloy, K., Mansouri, D.L., Martinez, E., Matejkova, P., Matlew, T., McCallum, S.K., McPherson, J.D., Mercado, C., Mercado, I.C., Metzger, M.L., Millin, A., Mlasevilevic, A., Morgan, M.B., Morris, S., Mundaka, M., Murray, D.D., Muzny, D.M., Nazarech, L.V., Ngo, D.N., Nguyen, H.T., Nguyen, N.B., Nguyen, P.O., Nwakoelameh, O.O., Obregon, M., Odeh, E.A., Okonko, F., Okunonu, G.O., Okunonu, K.C., Oryenewe, J., Parish, B.J., Parker, D.N., Parra, A.A., Pasternak, S., Patel, B.M., Patel, R.R., Paul, H.A., Perez, A., Perez, L.M., Perez, Y.Y., Pham, T.L., Player, E.J., Primus, E.L., Pu, L., Puzo, M., Purkiss, C., Qin, X., Quiroz, J.B., Rabata, D., Rachlin, E.K., Ren, Y., Richards, S., Rojas, A., Ruiz, S., Sabo, A., Santibanez, J., Savory, G.G., Scherer, S.E., Schneider, B.W., Sebasigiri, R., Sexon, M.M., Shen, H., Sleson, I., Sneed, A.J., Sodergren, E., Song, X., Sorrell, R.P., Svatek, A.E., Taylor, B.W., Taylor, T.R., Thelner, R., Thomas, N., Thorm, R.D., Thornton, R.D., Tong, M.Y., Trejos, Z.Y., Umanai, K., Vargo, C.B., Valtachli, S., Vega, R.A., Villana, D., Volkov, A., Walker, D.L., Wang, Q., Wang, S., Warren, J.T., Watt, J.E., Wei, X., Wheeler, D.A., White, C.S., Williams, Jr, R.L., Williams, A.C.,

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL

COMMENT

Williams, G.A., Williams, J.D., Wilson, K., Woodworth, J.R., Worley, K.C., Wright, R.A., Wu, J., Wu, W., Yakub, S., Yerrapragada, S., Yu, F., Yuan, D.T., Yuan, Y., Zhang, Y., Zhang, V., Zhang, Z., Zhou, J., Zhu, Y., Weinstock, G. and Gibbs, R.A.

Unpublished
Direct Submission
2 (bases 1 to 150038)
Worley, K.C.

Direct Submission
Submitted (01-FEB-2006) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center clone name: R3-4012A17
Sequencing Vector: pUC18
Cloning Vector: pBAC3.6
Chemistry: Dye-terminator Big Dye
Summary Statistics
Estimated insert size: sum-of-contigs estimation
Quality coverage: x in Q20 bases; sum-of-contigs estimation

NOTES: The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/atlantis/>). The BAC reads were identified through a deconvolution of an array containing pools of BACs. Due to the incorporated WGS reads, there may be contigs that consist entirely of whole genome shotgun sequence reads and the sequence may extend beyond the ends of the clone. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the features table.

NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 21261: contig of 21261 bp in length
21262 21271: gap of 10 bp
21272 31480: contig of 10209 bp in length
31481 31490: gap of 10 bp
31491 45811: contig of 14321 bp in length
45812 45821: gap of 10 bp
45822 62070: contig of 16249 bp in length
62071 62080: gap of 10 bp
62081 70393: contig of 8313 bp in length
70394 70493: gap of unknown length
70494 84032: contig of 13339 bp in length
84033 84042: gap of 10 bp
84043 106287: contig of 22245 bp in length
106288 106297: gap of 10 bp
106298 124992: contig of 18695 bp in length
124993 125002: gap of 10 bp
125003 130651: contig of 5649 bp in length
130652 130751: gap of unknown length
130752 138577: contig of 7826 bp in length
138578 138587: gap of 10 bp
138588 139792: contig of 1205 bp in length
139793 139892: gap of unknown length
139893 144331: contig of 4339 bp in length
144332 144333: gap of unknown length
144333 150038: contig of 5707 bp in length.

Location/Qualifiers
1. 150038
/organism="Strongylocentrotus purpuratus"
/mol_type="genomic DNA"
/db_xref="taxon:7668"

RESULT 15
AC1178783/c
LOCUS
DEFINITION
ACCESSION
AC1178783
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AC1178783 204539 bp DNA linear HTG 31-JAN-2006
Strongylocentrotus purpuratus clone R3-21G01, WORKING DRAFT
SEQUENCE, 31 unordered pieces.
AC1178783
HTG, HTGS_PHASE1, HTGS_DRAFT, HTGS_POOLED_CLONE, HTGS_ENRICHED.
Strongylocentrotus purpuratus
Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidae; Euechinoidae; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 204539)
Shen, Y., Abraham, K.K., Abulimiti, A., Adams, C.Q., Aduba, G.,
Allen, C.C., Alstrook, S.L., Anosike, U.S., Archer, P.M.,
Airdondo, H.H., Atchew, T., Bandaranaike, D.P., Bangura, L.,
Barton, S.R., Bell, A.V., Bell, S.N., Beraducci, A.R., Bicham, C.,
Blawalo, K., Blyth, P.R., Buhay, C.J., Canada, A., Cardenas, V.,
Carter, K., Chacko, J., Chandrabose, M.N., Chavez, A., Chavez, D.,
Chen, G., Chen, R., Chu, H., Clerc Blankenburg, K.P., Cockrell, R.,
Cooper, J.A., Coyle, M.D., Cree, A., Cuello, C.B., Curry, S.M., Dai, W.,
Dao, M.D., Davila, M., Davis, C., Davy-Carroll, L., Del Fierro, P.,
Demen, R., Denson, S., Ding, Y., Din, H.H., Donlin, J.E.,
Dugan-Bochs, S., Dunn, A.M., Durbin, K.J., Ebnong, V.E., Egan, A.,
Espinoza, V.C., Fa, M., Fernandez, S., Fernando, P.R., Ferrier, A.R.,
Flegg, N., Forbes, L.D., Fowler, R.G., Fu, Q., Fuh, E., Gabisi, R.A.,
Ganardhanan, M., Garner, J., Garcia, I.I., R.M., Garcia, A.M.,
Garcia, S.M., Garner, T.T., Ghouse, S., Gingras, M.,
Gonzalez-Garay, M.L., Guevara, W.V., Haaland, W.C., Heberlein, K.A.,
Hagens, B.J., Hall, O., Hamid, H., Hamilton, K.A., Hampton, O.A.,
Hartes, B.A., Harris, R.A., Havlik, P., Hawes, A.C., Hawkins, E.S.,
Haynes, S.J., Hemphill, L., Hernandez, J., Hines, S., Hirani, K.,
Hithens, M.E., Hodgson, A.V., Hogue, M.E., Holder, M., Hollins, B.,
Howell, L.L., Huik, S.W., Hume, J., Jackson, A., Jackson, L.R.,
Jacob, S.K., Jhangiani, S.N., Jiang, H., Johnson, B., Johnson, R.,
Joishi, V., Joy, C., Kalkal, P.B., Kalafatis, K.J., Kalu, J.B., Kang, Y.,
Keeler, J., Khan, Z.M., Kidwai, S., King, L.M., Kisman, H., Kovac, C.L.,
Kowals, A.N., Kowis, C.R., Lago, L.A., Lago, M.T., Lai, C., Lare, F.,
Le, T.T., Lee, S.L., Lee, T.W., Legall, I.I., F.H., Lemon, S.J.,
Lewis, L.R., Li, B., Li, Y., Li, Z., Linell, M.A., Liu, J., Liu, W.,
Liu, Y., Liu, Y., Lijang, D., London, P., Lopez, J., Lorensen, L.M.,
Lozdo, R.J., Luc, T., Madu, R.C., Maheshwari, M., Maheshwari, L.M.,
Malloy, K., Mansouri, D.L., Martinez, E., Matejkova, P., Mathew, T.,
McCauley, S.K., McPherson, J.D., Mercado, C., Mercado, I.C.,
Meeker, M.L., Millin, A., Milosavljevic, A., Morgan, M.B., Morris, S.,
Mundasa, M., Murray, D.D., Muzny, D.M., Nazareth, L.V., Ngo, D.N.,
Nguyen, H.T., Nguyen, N.B., Nguyen, P.Q., Nwaokeme, O.O.,
Obregon, M., Odeh, E.A., Okonko, F., Okunolu, G.O., Okunolu, K.C.,
Oyeneke, J., Parish, B.J., Parker, D.N., Parra, A.A., Paternak, S.,
Patel, B.M., Patel, R.R., Paul, H.A., Perez, A., Perez, L.M.,
Perez, Y.Y., Pham, T.L., Player, E.J., Primus, B.L., Pu, L., Puzo, M.,
Purkiss, C., Qin, X., Quiroz, J.B., Rabata, D., Rachlin, E.K., Ren, Y.,
Richards, S., Rojas, A., Ruiz, S., Sabo, A., Santibanez, J.,
Savory, G.G., Scherer, S.E., Schneider, B.W., Sebasigiri, R.,
Sexon, M.M., Shen, H., Sisson, I., Sneed, A.J., Sodergren, E., Song, X.,
Soralle, R.P., Svatek, A.F., Taylor, E.W., Taylor, T.R., Theuber, R.,
Thomas, N., Thorm, R.D., Thornton, R.D., Tong, M.Y., Trejos, Z.Y.,
Umanan, K., Varghese, C.E., Valtachil, S., Vega, R.A., Villalana, D.,
Velkov, A., Walker, D.L., Wang, Q., Wang, S., Warren, J.T., Watt, J.E.,
Wei, X., Wheeler, D.A., White, C.S., Williams, J.R., Williams, A.C.,
Williams, G.A., Williams, J.D., Wilson, K., Woodworth, J.R.,
Worley, K.C., Wright, R.A., Wu, J., Wu, W., Yakub, S., Yerragada, S.,
Yu, F., Yuan, D.T., Yuan, Y., Zhang, J., Zhang, L., Zhang, Z., Zhou, J.,
Zhu, Y., Weinstein, G. and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 204539)
Unpublished
Direct Submission
Worley, K.C.
Submitted (31-JAN-2006) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center clone name: R3-21G01
Sequencing Vector: pUC18
Cloning Vector: pBACe3.6
Chemistry: Dye-terminator Big Dye
Summary Statistics
Estimated insert size: i sum-of-contigs estimation
Quality coverage: x in Q20 bases: sum-of-contigs estimation

NOTE: The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). The BAC reads were identified through a deconvolution of an array containing pools of BACs. Due to the incorporated WGS reads, there may be contigs that consist entirely of whole genome shotgun sequence reads and the sequence may extend beyond the ends of the clone. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the features table.

NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 14737: contig of 14737 bp in length
14738 15709: gap of 972 bp
15710 18398: contig of 2689 bp in length
18399 18408: gap of 10 bp
18409 23389: contig of 4981 bp in length
23390 23399: gap of 10 bp
23400 23647: contig of 248 bp in length
23648 23657: gap of 10 bp
23658 34000: contig of 10343 bp in length
34001 34010: gap of 10 bp
34011 39118: contig of 5108 bp in length
39119 39128: gap of 10 bp
39129 47794: contig of 8666 bp in length
47795 47804: gap of 10 bp
47805 48904: contig of 1100 bp in length
48905 49004: gap of unknown length
49005 51218: contig of 2214 bp in length
51219 51318: gap of unknown length
51319 63376: contig of 12058 bp in length
63377 63386: gap of 10 bp
63387 67353: contig of 3967 bp in length
67354 67363: gap of 10 bp
67364 77092: contig of 9729 bp in length
77093 77102: gap of 10 bp
77103 92743: contig of 15641 bp in length
92744 92753: gap of 10 bp
92754 103572: contig of 10819 bp in length
103573 103672: gap of unknown length
103673 105208: contig of 1536 bp in length
105209 105218: gap of 10 bp
105219 106802: contig of 1584 bp in length
106803 106802: gap of unknown length
106803 108811: contig of 1909 bp in length
108812 108821: gap of 10 bp
108822 139372: contig of 30551 bp in length
139373 139472: gap of unknown length
139473 140424: contig of 952 bp in length
140425 141670: gap of 1246 bp in length
141671 143006: contig of 1336 bp in length
143007 144509: gap of unknown length
144510 144609: gap of unknown length

*	144610	145235	contig of 626 bp in length
*	145226	145335	gap of unknown length
*	145336	146750	contig of 1415 bp in length
*	146751	146850	gap of unknown length
*	146851	147639	contig of 789 bp in length
*	147640	147739	gap of unknown length
*	147740	148988	contig of 1249 bp in length
*	148989	149088	gap of unknown length
*	149089	149927	contig of 839 bp in length
*	149928	150027	gap of unknown length
*	150028	154658	contig of 4631 bp in length
*	154659	154758	gap of unknown length
*	154759	156594	contig of 1836 bp in length
*	156595	156694	gap of unknown length
*	156695	157836	contig of 1142 bp in length
*	157837	157936	gap of unknown length
*	157937	167699	contig of 9763 bp in length
*	167700	167799	gap of unknown length
*	167799	204539	contig of 36740 bp in length

FEATURES	Location/Qualifiers
source	1. .204539

```

/organism="Strongylocentrotus purpuratus"
/mol_type="genomic DNA"
/db_xref="taxon:7668"
/clone="RR3-21601"

1. 14737
/misc_feature
/notes="assembly_name:Contig28"
14738. .15709
/misc_feature
/notes="assembly_name:gap"
14738. .15709
/misc_feature
/estimated_length=972
15710. .18398
/misc_feature
/notes="assembly_name:Contig16"
18399. .18408
/misc_feature
/notes="assembly_name:gap"
18399. .18408
/misc_feature
/estimated_length=10
18409. .23389
/misc_feature
/notes="assembly_name:Contig21"
23390. .23399
/misc_feature
/notes="assembly_name:gap"
23390. .23399
/misc_feature
/estimated_length=10
23400. .23647
/misc_feature
/notes="assembly_name:Contig1"
23648. .23657
/misc_feature
/notes="assembly_name:gap"
23648. .23657
/misc_feature
/estimated_length=10
23658. .34000
/misc_feature
/notes="assembly_name:Contig24"
34001. .34010
/misc_feature
/notes="assembly_name:gap"
34001. .34010
/misc_feature
/estimated_length=10
34011. .39118
/misc_feature

```

Query Match	5.0%	Score 94.6	DB 12	Length
Best Local Similarity	44.0%	Pred. No. 1.4e-16		
Matches 455	Conservative	0	Mismatches 574	Indels

[illegible]

Oy	278	CCGCAAGCGTTGAATCAGCTAACAACAGCTGCCAAAGTTTAACTCCTGGAAACCTGTTC	337
Dd	133413	TBACTTAAGACAAAGAACGAGAGAGAGAAGTAGTGATAAAGACGTAAAAGAAAGAAAGAAAG	133355
Oy	338	CTCAAGAAACCAACCGATGCMAATGGGAATGTCACGTTCATTCTAAAAAACAATG	397
Dd	133353	AAGCGTAAAAAGAAAGAAAGTAGTAGAAAGAAAGTAAAAAGAAAGAAAGAAAGTATATAAAAAGAAAG	133294
Oy	358	GTAAGAGATGCAGTGTATACCATTAAAGAAACCAAAGCGGTGTATGTTCTGTCCGA	457
Dd	133293	AAAAAGAGACAGACAGACAGAAAGAAAGAAAGTAAAGACAGAGTAGTAGTAAAAAGAAAGTAAAG	133233
Oy	458	ATATGCTGGTGGCGTTCCCACTTTACGAATGATCAAGCAACAGATGCTTCTATTAAT	517
Dd	133233	AAGAAGAAAGACAGATPAAAAAGAAAGTAGTAAAAAGATPAAAAAGAAAGAAAGAAAG	133174
Oy	518	ATGGAACAAGAAATTAGCCGTGTTCTCATTTATTCCTAAAAATGTGATGCCAATGATG	577
Dd	133173	AAGAAGAAAGTAAAGACAGAGTAGTAAAAAGAAAGACAGACAGAGTAGTAAAAAGAAAGTAG	133113

578 GTAGTTACATGTGAAAAAGTAGCACTGCTGAAAATGAAGATTAAATGGCGCAGAAT 637

Db	133113	TAAAGAAAGTAAAGAGAGAAAGAAAGCGTAAAGAGAAAGTGTAAAGAAAGTAAAG	133054
Qy	638	TTGTTATTTCTAAAGCCGAGGCTCACCGGACAGTAAATATATCCAGAGTCAAG	697
Db	133053	AAG-----AAGAGAGAGAGAGAGAGCGTAAAGAGAGAGAGTGTAAAG	133000
Qy	698	ATGATTTATATCATGGACAAACGATAAAGAAACAGCAAAACGCTTTATTACTGGGAAA	757
Db	132959	AAGAAGAGTTATTAAGAAAAAGAAAAAGAAAGAAAGAAAGAAAGTAAAGAG	132940
Qy	758	GTTATGAAGTTGGCCAAAATGATTTACAGAAAGCAGAGATGGACGGGGAATTAAAG	817
Db	132939	AAGTGTAAAGAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGTAAAGAG	132880
Qy	818	TTAAAGATCTTGAGGTGGTTCGTATATTTAGAGAGTAAAGCTCCAAATATGACAG	877
Db	132879	TAAAGAGAGAGAGAGAGAGAGAGTAAAGAGAGAGAGAGAGTGTAAAGAGTAAAGAGAG	132820
Qy	878	AATTAATTGAAAAATCAAACAAAAACCACTTTACATTTGAGCAACATCAACACCTG	937
Db	132819	AAGTGTAAAGAGAGAGAG	132760
Qy	938	TTGAAGAAAAACGTCAAAAAATGATATCTCTTAAGATTATTAAGAAACCAACGCTTAGATG	997
Db	132759	TTATTAAGAAAAAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGTAA	132700
Qy	998	GTAAGATGTGGCAATTGGCCGAAAAAATTAATATCAAAATTTCTGTAAATATTCATTGG	1057
Db	132639	AAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGTAAAGAGAG	132640
Qy	1058	GGATTTGCAGCAAGAGAGCGAGCGTAAATTAATAGTCAAAATTCAAATTTAGTTGATTAAC	1117
Db	132639	AAGAGAGAGAGAGAGAGAGTGTAAAGAGAGAGAGAGAGTGTAAAGAGAGAGAGAGTAA	132580

QY	1118	ATGATGCAGCCTTAA	1132
Db	132579	AAGAAGAACTAGTAA	132565

Search completed: July 5, 2006, 20:23:46
Job time : 7258 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW nucleic - nucleic search, using sw model

Run on: July 5, 2006, 18:22:58 ; Search time 6165 Seconds
(without alignments)
17061.519 Million cell updates/sec

Title: US-10-661-809A-12

Perfect score: 1881
Sequence: 1 atgaagcaatcaaaaaagt.....gacgtagaagaagaatgct 1881

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_esc1:*
2: gb_esc3:*
3: gb_esc4:*
4: gb_esc5:*
5: gb_esc6:*
6: gb_esc7:*
7: gb_esc8:*
8: gb_esc9:*
9: gb_esc10:*
10: gb_esc11:*
11: gb_esc12:*
12: gb_esc13:*
13: gb_esc14:*
14: gb_esc15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	496.4	26.4	500	11	B07882 YX1027T PBI
2	492.6	26.2	500	11	B07880 YX988T PBI
3	492	26.2	500	11	B07866 YX1067T PBI
4	434.8	23.1	450	11	B07865 YX10673 PBI
5	430.2	22.9	450	11	B07881 YX10273 PBI
6	380	20.2	400	11	B07889 YX988T PBI
7	91.2	4.8	908	11	AZ548467 ENTEK30TR
8	88.8	4.7	843	11	AZ551618 ENTEK30TR
9	79	4.2	1084	11	AL049509 T7 end of
10	78.2	4.2	1811	12	CG53732 P048-4-G0
11	78	4.1	877	11	AZ531291 ENTBQ34TR
12	77.6	4.1	1416	9	DN711652 ENTBQ34TR
13	77.2	4.1	912	11	AZ551092 ENTBQ34TR
14	77	4.1	879	14	CNS01JRG
15	76.4	4.1	993	7	BH898130
16	76.4	4.1	1496	12	CU073855 CH216-130
17	76	4.0	1104	1	AJ925669
18	75.8	4.0	1981	12	CU082000 CH216-165
19	75.6	4.0	1255	14	AG435644 Mus muscu

20	75	4.0	1210	12	CG751315	CG751315 P045-4-B0
21	75	4.0	1687	1	AM090701	AM090701 AGENCOURT
22	74.4	4.0	1300	2	BM468018	BM468018 AGENCOURT
23	74.2	3.9	1053	3	BUS08694	BUS08694 AGENCOURT
24	73.8	3.9	868	11	BH161951	BH161951 ENTBQ34TR
25	73.8	3.9	1348	9	DN721897	DN721897 CNB140-H1
26	73.2	3.9	1896	10	DV780693	DV780693 Hw PAT 11
27	73	3.9	1346	14	AG429454	AG429454 Mus muscu
28	72.8	3.9	949	4	CA46506	CA46506 AGENCOURT
29	72.8	3.9	1687	12	CU082679	CU082679 CH216-169
30	72.6	3.9	1101	14	CNS0039G	AL063921 Drosophila
31	72.6	3.9	1260	13	CU491610	CU491610 SAIT 559
32	72.6	3.9	1753	9	DR131208	DR131208 49217594
33	72.4	3.8	905	11	AZ550256	AZ550256 ENTEK30TR
34	72.4	3.8	1058	4	CD048666	CD048666 AGENCOURT
35	72.2	3.8	1493	12	CU078589	CU078589 CH216-151
36	72	3.8	964	13	CM957121	CM957121 TCBS2.1C
37	72	3.8	1114	9	DN656077	DN656077 CEC21-A11
38	72	3.8	1594	14	CU038406	CU038406 CH216-46A
39	72	3.8	1843	12	AG435185	AG435185 Mus muscu
40	71.8	3.8	1256	12	CU119201	CU119201 ISB1-76J1
41	71.8	3.8	1459	12	CU040025	CU040025 CH216-49A
42	71.8	3.8	1566	12	CG757757	CG757757 P053-1-D0
43	71.6	3.8	1578	10	DV791374	DV791374 HM 101n 1
44	71.6	3.8	975	3	BU960792	BU960792 AGENCOURT
45	71.6	3.8	1059	14	CNS0022B	AL097133 Drosophila

ALIGNMENTS

RESULT 1
B07882/c
LOCUS B07882 500 bp DNA GSS 13-MAR-1998
DEFINITION YX1027T plasmid expression library of genomic DNA from OGI-R
Enterococcus faecalis genomic clone YX102 similar to hypothetical
protein 2 of Lactobacillus leichmannii (X81869), genomic survey
sequence.

ACCESSION B07882
VERSION B07882.1 GI:2058762
KEYWORDS GSS.
SOURCE Enterococcus faecalis
ORGANISM Enterococcus faecalis
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;
Enterococcus.
REFERENCE 1 (bases 1 to 500)
AUTHORS Xu, Y., Jiang, L., Murray, B. E. and Weinstein, G. M.
TITLE Enterococcus faecalis antigens in human infections
JOURNAL Infect. Immun. 65 (10), 4207-4215 (1997)
PUBMED 9317028
COMMENT Contact: Weinstein, G M
Department of Biochemistry and Molecular Biology
University of Texas Medical School
6431 Fannin Street, Houston, TX 77030, USA
Tel: 713 500 6083
Fax: 713 500 0652
Email: georgew@utmsi.med.utmc.edu
Seq primer: plasmid SK (-) c7
Clas: Shotgun
High quality sequence stop: 500.
Location/Qualifiers
1..500
/organism="Enterococcus faecalis"
/mol_type="genomic DNA"
/strain="OGIRP"
/db_xref="taxon:1351"
/clone="YX102"
/lab_host="XLRblue MRF"
/clone_lib="plasmid expression library of genomic DNA
from OGI-R"
/note="Vector: plasmid SK (-), Site 1: HindIII, Site 2:
SCOR; genomic DNA from an Enterococcus faecalis strain
OGIRP was partially digested with SauAI, sized

FEATURES

source

fractionated and ligated into the BamH site of cosmid vectors pLARX and pBelobAC11. After in vitro packaging and plating, the cosmid clones were screened with antisera from 4 enterococcal endocarditis patients and a rabbit immunized with surface proteins from an E. faecalis isolate. Cosmid DNA from immunopositive cosmid clones was treated with Dnae I, the ends were filled with the Klenow fragment, and then ligated into the EcoRV site in pBluescript SK(-). The ligation mixture was transformed into XLBlue MRF⁺, and the transformants were screened with antisera. The immunopositive clones were selected for DNA sequencing."

ORIGIN

```
Query Match      26.4%; Score 496.4; DB 11; Length 500;
Best Local Similarity 99.4%; Pred. No. 6.2e-106;
Matches 497; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1319 AAAAAGCAGATCCTACGAAAGGCTTTAAATAAGAGCGAATGTGATAAGGTCATACCG 1378
DB 500 AAAAGCAGATCCTACGAAAGGCTTTAAATAAGAGCGAATGTGATAAGGTCATACCG 441
QY 1379 ACGACCAAAACACCACTGTTGAAGTGTGACAGGTGGGAAACGTTTCATTAAAGTCG 1438
DB 440 ACGACCAAAACACCACTGTTGAAGTGTGACAGGTGGGAAACGTTTCATTAAAGTCG 381
QY 1439 ATGGGAGTGTACACGACACAAAGCCTTGCGCGAGCTTCCTTGTCGCGCGATCAAA 1498
DB 380 ATGGGAGTGTACACGACACAAAGCCTTGCGCGAGCTTCCTTGTCGCGCGATCAAA 321
QY 1499 ACAGCGACACACCAAAATTATTTGAAATGATGAAACAAAGCAAGCAAGCACTTGGGTGA 1558
DB 320 ACAGCGACACACCAAAATTATTTGAAATGATGAAACAAAGCAAGCAAGCACTTGGGTGA 261
QY 1559 AAACAAAGCTGAAGCACTTTTACAAACGCGTGTGATGATGATTGATATACAG 1618
DB 260 AAACAAAGCTGAAGCACTTTTACAAACGCGTGTGATGATGATATACAG 201
QY 1619 GGCCTAAATAGGGTACCTTATTTAGAGAAAGCTAGCTCTGATGATTAATGCTTGT 1678
DB 200 GGCCTAAATAGGGTACCTTATTTAGAGAAAGCTAGCTCTGATGATTAATGCTTGT 141
QY 1679 TAACAATCGATGTAATTTGTGTCATGAAACAATCATATGAGCAACAAGAAACCTAG 1738
DB 140 TAACAATCGATGTAATTTGTGTCATGAAACAATCATATGAGCAACAAGAAACCTAG 81
QY 1739 TTTCCACGAAAAAGTACCAAAACAAGATGCTTACCTTCAACAGGTGGCAAG 1798
DB 80 TTTCCACGAAAAAGTACCAAAACAAGATGCTTACCTTCAACAGGTGGCAAG 21
QY 1799 GAATCTAGCTTACTTAGGA 1818
DB 20 GAATCTAGCTTACTTAGGA 1

RESULT 2      500 bp      DNA      linear      GSS 13-MAR-1998
B07890/c      B07890/c
LOCUS          B07890/c
DEFINITION    YX98T7 pBluescript expression library of genomic DNA from OG1RF
              Enterococcus faecalis genomic clone YX98 similar to hypothetical
              protein 2 of Lactobacillus leichmannii (X81869), genomic survey
              sequence.
ACCESSION     B07890
VERSION       B07890.1 GI:2058780
KEYWORDS      GSS.
SOURCE        Enterococcus faecalis
              Enterococcus faecalis
              Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;
              Enterococcus.
REFERENCE     1 (bases 1 to 500)
              Xu, Y., Jiang, L., Murray, B.E. and Weinstein, G.M.
              Enterococcus faecalis antigens in human infections
              Infect. Immun. 65 (10), 4207-4215 (1997)
```

PUBMED

9317028
Contact: Weinstein, GM
Department of Biochemistry and Molecular Biology
University of Texas Medical School
6431 Fannin Street, Houston, TX 77030, USA
Tel: 713 500 6083
Fax: 713 500 0652
Email: georgew@utmsf.med.utmc.edu
Seq primer: pBluescript SK (-) t7
Class: shotgun
High quality sequence Etop: 500.
Location/Qualifiers

FEATURES

source

```
1..500
/organism="Enterococcus faecalis"
/mol_type="genomic DNA"
/strain="OG1RF"
/db_xref="taxon:1351"
/clone="YX98"
/lab_host="XLBlue MRF"
/clone_1ib="pBluescript expression library of genomic DNA
from OG1RF"
/note="vector: pBluescript SK(-); Site_1: HindIII; Site_2:
EcoRI; Genomic DNA from an Enterococcus faecalis strain
OG1RF was partially digested with Sau3AI, sized
fractionated and ligated into the BamH site of cosmid
vectors pLARX and pBelobAC11. After in vitro packaging
and plating, the cosmid clones were screened with antisera
from 4 enterococcal endocarditis patients and a rabbit
immunized with surface proteins from an E. faecalis
isolate. Cosmid DNA from immunopositive cosmid clones was
treated with Dnae I, the ends were filled with the Klenow
fragment, and then ligated into the EcoRV site in
pBluescript SK(-). The ligation mixture was transformed
into XLBlue MRF+, and the transformants were screened
with antisera. The immunopositive clones were selected for
DNA sequencing."
```

ORIGIN

```
Query Match      26.2%; Score 492.6; DB 11; Length 500;
Best Local Similarity 99.0%; Pred. No. 4.9e-105;
Matches 495; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1190 TGATTCCTCTGAAAATATATCAAGTGAACAAGAAATGCTTCACTGTCGCGTTA 1249
DB 500 TGATTCCTCTGAAAATATATCAAGTGAACAAGAAATGCTTCACTGTCGCGTTA 441
QY 1250 ATCCAGCGTATATTCCTAGCGTAAACACGAGGCGGCACTAAATTCCTTACTTATGC 1309
DB 440 ATCCAGCGTATATTCCTAGCGTAAACACGAGGCGGCACTAAATTCCTTACTTATGC 381
QY 1310 ATTTAAATGAATAAGCAGATCCTACGAAAGCTTTAAATAAGGCGAATGTTGATACG 1369
DB 380 ATTTAAATGAATAAGCAGATCCTACGAAAGCTTTAAATAAGGCGAATGTTGATACG 321
QY 1370 GTCATACGACGACCAACCAACCACTGTTGAAGTGTGACAGGTGGAAACGTTTCA 1429
DB 320 GTCATACGACGACCAACCAACCACTGTTGAAGTGTGACAGGTGGAAACGTTTCA 261
QY 1430 TTTAAAGTGATGCGCATGTGACAGCAACAAGCTTGGCGGAGCTTCTTTCGCTGC 1489
DB 260 TTTAAAGTGATGCGCATGTGACAGCAACAAGCTTGGCGGAGCTTCTTTCGCTGC 201
QY 1490 GTGATCAAAACAGGACACAGCAATTTATTTGAAATGATGATGAACAAGAAAGACGAA 1549
DB 200 GTGATCAAAACAGGACACAGCAATTTATTTGAAATGATGATGAACAAGAAAGACGAA 141
QY 1550 CTTGGGTGAAAACAAAGCTGAAGCACTTTTACAAACAAGCTGATGATTAAGTTG 1609
DB 140 CTTGGGTGAAAACAAAGCTGAAGCACTTTTACAAACAAGCTGATGATTAAGTTG 81
QY 1610 ATATCAAGGGCTTAAATACGCTATATTATTTGAAAGAAACGTAGTCTCTGATGATT 1669
DB 80 ATATCAAGGGCTTAAATACGCTATATTATTTGAAAGAAACGTAGTCTCTGATGATT 21
```

QY 1670 ATGCTGTGTAACAATCGG 1689
DB 20 ATGCTGTGTAACAATCGG 1

RESULT 3
B07866/c
LOCUS
DEFINITION B07866 500 bp DNA linear GSS 13-MAR-1998
YX106T3 pBluescript expression library of genomic DNA from OG1RF
Enterococcus faecalis clone YX106 similar to hypothetical
protein 2 of Lactobacillus leichmannii (X81869), genomic survey
sequence.

ACCESSION B07866
VERSION B07866.1 GI:2058766
KEYWORDS GSS.
SOURCE Enterococcus faecalis
ORGANISM Enterococcus faecalis
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;
Enterococcus.

REFERENCE 1 (bases 1 to 500)
AUTHORS Xu, Y., Jiang, L., Murray, B.E. and Weinstein, G.M.
TITLE Enterococcus faecalis antigens in human infections
JOURNAL Infect. Immun. 65 (10), 4207-4215 (1997)
PUBMED 9317028

COMMENT Contact: Weinstein, GM
Department of Biochemistry and Molecular Biology
University of Texas Medical School
6431 Fannin Street, Houston, TX 77030, USA
Tel: 713 500 6083
Fax: 713 500 0652
Email: georgew@utmsf.med.utmc.edu
Seg primer: pBluescript SK (-) c7
Class: shotgun
High quality sequence stop: 500.
Location/Qualifiers
1..500
/organism="Enterococcus faecalis"
/mol_type="genomic DNA"
/strain="OG1RF"
/db_xref="taxon:1351"
/clone="YX106"
/lab_host="XL1Blue MRF"
/clone_1lb="pBluescript expression library of genomic DNA
from OG1RF"
/note="Vector: pBluescript SK(-); Site 1: HindIII; Site 2:
EcoRI; Genomic DNA from an Enterococcus faecalis strain
OG1RF was partially digested with Sau3AI, sized
fractionated and ligated into the BamHI site of cosmid
vectors pLAFx and pReloBAC11. After in vitro packaging
and plating, the cosmid clones were screened with antisera
from 4 enterococcal endocarditis patients and a rabbit
immunized with surface proteins from an E. faecalis
isolate. Cosmid DNA from immunopositive cosmid clones was
treated with Dnae I, the ends were filled with the Klenow
fragment, and then ligated into the EcoRV site in
pBluescript SK(-). The ligation mixture was transformed
into XL1Blue MRF, and the transformants were screened
with antisera. The immunopositive clones were selected for
DNA sequencing."

ORIGIN
Query Match 26.2%; Score 492; DB 11; Length 500;
Best Local Similarity 99.0%; Pred. No. 6.8e-105;
Matches 495; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 997 GGTAAAGATGTCGCAATTCGCAAAAATTAATATCAATTCGTAATATTCGATG 1056
DB 500 GGTAAAGATGTCGCAATTCGCAAAAATTAATATCAATTCGTAATATTCGATG 441

QY 1057 GGGATTGCAGACAAAGAGCGCATTAATAATACGTAATTCATTTAGTTGATAA 1116
DB 440 GGGATTGCAGACAAAGAGCGCATTAATAATACGTAATTCATTTAGTTGATAA 381

QY 1117 CATGATGACGCTTAACCTTTGATTAACGATCTTCGAGAGTATGCTTATAT 1176
DB 380 CATGATGACGCTTAACCTTTGATTAACGATCTTCGAGAGTATGCTTATAT 321

QY 1177 GATGGGATATACGATGATTCCTCTGTAATAATTAATCAAGTATGAAACAAATGCTTC 1236
DB 320 GATGGGATATACGATGATTCCTCTGTAATAATTAATCAAGTATGAAACAAATGCTTC 261

QY 1237 ACTGTGCGCGTTAATTCACGATATATCTTACGCTTACACGACGCGGACACTTAAATTC 1296
DB 260 ACTGTGCGCGTTAATTCACGATATATCTTACGCTTACACGACGCGGACACTTAAATTC 201

QY 1297 GTTACTTATGATTAATGAAAGAGATCCATGAGAAAGGCTTTAAATGAGGCG 1356
DB 200 GTTACTTATGATTAATGAAAGAGATCCATGAGAAAGGCTTTAAATGAGGCG 141

QY 1357 AATGTTGATTAACGATCATACGACGACCAACCAACCACTTTGAAAGTTGACAGGT 1416
DB 140 AATGTTGATTAACGATCATACGACGACCAACCAACCACTTTGAAAGTTGACAGGT 81

QY 1417 GGGAAACGTTTCATTAAGTCATGCGCATGTGACGACGACCAACGCTTGGCGGAGCT 1476
DB 80 GGGAAACGTTTCATTAAGTCATGCGCATGTGACGACGACCAACGCTTGGCGGAGCT 21

QY 1477 TCCTTGTGTCGCGTGATCA 1496
DB 20 TCCTTGTGTCGCGTGATCA 1

RESULT 4
B07865
LOCUS
DEFINITION YX106T3 pBluescript expression library of genomic DNA from OG1RF
Enterococcus faecalis clone YX106, genomic survey sequence.

ACCESSION B07865
VERSION B07865.1 GI:2058765
KEYWORDS GSS.
SOURCE Enterococcus faecalis
ORGANISM Enterococcus faecalis
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;
Enterococcus.

REFERENCE 1 (bases 1 to 450)
AUTHORS Xu, Y., Jiang, L., Murray, B.E. and Weinstein, G.M.
TITLE Enterococcus faecalis antigens in human infections
JOURNAL Infect. Immun. 65 (10), 4207-4215 (1997)
PUBMED 9317028

COMMENT Contact: Weinstein, GM
Department of Biochemistry and Molecular Biology
University of Texas Medical School
6431 Fannin Street, Houston, TX 77030, USA
Tel: 713 500 6083
Fax: 713 500 0652
Email: georgew@utmsf.med.utmc.edu
Seg primer: pBluescript SK (-) c3
Class: shotgun
High quality sequence stop: 450.
Location/Qualifiers
1..450
/organism="Enterococcus faecalis"
/mol_type="genomic DNA"
/strain="OG1RF"
/db_xref="taxon:1351"
/clone="YX106"
/lab_host="XL1Blue MRF"
/clone_1lb="pBluescript expression library of genomic DNA
from OG1RF"
/note="Vector: pBluescript SK(-); Site 1: HindIII; Site 2:
EcoRI; Genomic DNA from an Enterococcus faecalis strain
OG1RF was partially digested with Sau3AI, sized
fractionated and ligated into the BamHI site of cosmid
vectors pLAFx and pReloBAC11. After in vitro packaging
and plating, the cosmid clones were screened with antisera

from 4 enterococcal endocarditis patients and a rabbit immunized with surface proteins from an E. faecalis isolate. Cosmid DNA from immunopositive cosmid clones was treated with Dnae I, the ends were filled with the Klenow fragment, and then ligated into the EcoRV site in pBluescript SK(-). The ligation mixture was transformed into XL1Blue MRF, and the transformants were screened with antisera. The immunopositive clones were selected for DNA sequencing."

ORIGIN

Query Match 23.1%; Score 434.8; DB 11; Length 450;
Best Local Similarity 99.1%; Pred. No. 2e-91;
Matches 447; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

QY 725 AAGAACAAAGCAAAAGCGTTTATTACTGGGAAAAGTTATGAAATTTGGCGAAATGATTTCA 784
DB 1 AAGAACAAAGCAAAAGCGTTTATTACTGGGAAAAGTTATGAAATTTGGCGAAATGATTTCA 60
QY 785 CAGAGCAGAGAAATGGAACGGGAGATTAAAGTTTAAATCTTGAGGTGGTGTATATA 844
DB 61 CAGAGCAGAGAAATGGAACGGGAGATTAAAGTTTAAATCTTGAGGTGGTGTATATA 120
QY 845 TTTTGAAGAAAGTAAAGCTCCAAATTAATGCAATTAATGAAAATCAACAAAACAC 904
DB 121 TTTTGAAGAAAGTAAAGCTCCAAATTAATGCAATTAATGAAAATCAACAAAACAC 180
QY 905 CATTTACAAATTGAAACAAACATCAACACCTGTTGAAAAAACAGTCAAAATGATACCT 964
DB 181 CATTTACAAATTGAAACAAACATCAACACCTGTTGAAAAAACAGTCAAAATGATACCT 240
QY 965 CTAAGTGTATTAACCAACACCAAGCTGATGATGTTAAAGATGGCAATGGCGAAAAA 1024
DB 241 CTAAGTGTATTAACCAACACCAAGCTGATGATGTTAAAGATGGCGAAAAA 300
QY 1025 TTTAAATATCAAAATTTCTGTAATATTCATTTGGGGAATTCAGACAAAGACGAGCTA 1084
DB 301 TTTAAATATCAAAATTTCTGTAATATTCATTTGGGGAATTCAGACAAAGACGAGCTA 359
QY 1085 ATTAATAGTCGAATTCATTTAGTTGATTAACATGATGACAGCTTAATCTTTTGATACG 1144
DB 360 ATTAATAGTCGAATTCATTTAGTTGATTAACATGATGACAGCTTAATCTTTTGATACG 419
QY 1145 TGACTTCTGAGAGATGATGCTTATGCGTTATA 1175
DB 420 TGACTTCTGAGAGATGATGCTTATGCGTTATA 450

```

RESULT 5

B07881 450 bp DNA linear GSS 13-MAR-1998
LOCUS YX102T3 pBluescript expression library of genomic DNA from OG1RF
DEFINITION Enterococcus faecalis genomic clone YX102, genomic survey sequence.
ACCESSION B07881
VERSION B07881.1 GI:2058761
KEYWORDS GSS.
SOURCE Enterococcus faecalis
ORGANISM Enterococcus faecalis
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;
Enterococcus.
1 (bases 1 to 450)
Xu, Y., Jiang, L., Murray, B.E. and Weinstock, G.M.
Enterococcus faecalis antigens in human infections
Infect. Immun. 65 (10), 4207-4215 (1997)
9317028
Contact: Weinstock, GM
Department of Biochemistry and Molecular Biology
University of Texas Medical School
6431 Fannin Street, Houston, TX 77030, USA
Tel: 713 500 6083
Fax: 713 500 0652
Email: georgiew@utmsg.med.utah.edu
Seq primer: pBluescript SK(-) c3

Class: Shotgun
High quality sequence strop: 450.
Location/Qualifiers

FEATURES

source

1..450
/organism="Enterococcus faecalis"
/mol_type="genomic DNA"
/strain="OG1RF"
/db_xref="taxon:1351"
/clone="YX102"
/lab_host="XL1Blue MRF"
/clone_11b="pBluescript expression library of genomic DNA from OG1RF"
/note="Vector: pBluescript SK(-); Site 1: HindIII; Site 2: EcoRI; Genomic DNA from an Enterococcus faecalis strain OG1RF was partially digested with Sau3AI, sized fractionated and ligated into the BamHI site of cosmid vectors pLAFx and pReloBAC11. After in vitro packaging and plating, the cosmid clones were screened with antisera from 4 enterococcal endocarditis patients and a rabbit immunized with surface proteins from an E. faecalis isolate. Cosmid DNA from immunopositive cosmid clones was treated with Dnae I, the ends were filled with the Klenow fragment, and then ligated into the EcoRV site in pBluescript SK(-). The ligation mixture was transformed into XL1Blue MRF, and the transformants were screened with antisera. The immunopositive clones were selected for DNA sequencing."

ORIGIN

Query Match 22.9%; Score 430.2; DB 11; Length 450;
Best Local Similarity 98.2%; Pred. No. 2.5e-90;
Matches 435; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 1070 AAGAGGGAGCGCTTAATAATCTGCAATTCATATTGATGATAAATGACGCT 1129
DB 7 AAGAGGGAGCGCTTAATAATCTGCAATTCATATTGATGATAAATGACGCT 66
QY 1130 TAACTTTGATTAACGCTGACTTCTGAGAGATAGCTTAATGCGTTATATGATGAGGATACAG 1189
DB 67 TAACTTTGATTAACGCTGACTTCTGAGAGATAGCTTAATGCGTTATATGATGAGGATACAG 126
QY 1190 TGATTCCTCTGAAATTAATCAAGTACTGAACAAAGAGCTTCACTGCGCGTTA 1249
DB 127 TGATTCCTCTGAAATTAATCAAGTACTGAACAAAGAGCTTCACTGCGCGTTA 186
QY 1250 ATCCAGCGTATATTCCTACGCTTAACACGAGGCGACACTAAATTCGTTACTTTATGC 1309
DB 187 ATCCAGCGTATATTCCTACGCTTAACACGAGGCGACACTAAATTCGTTACTTTATGC 246
QY 1310 ATTTAAATGAAAAGACAGATCCTACGAAAGGCTTTAAATGAGGCGAAATGTTGATACG 1369
DB 247 ATTTAAATGAAAAGACAGATCCTACGAAAGGCTTTAAATGAGGCGAAATGTTGATACG 306
QY 1370 GTCATACCGACGACCAACCAACCACTGTTAAAGTTGTGACAGGTGGAAAAGCTTTCA 1429
DB 307 GTCATACCGACGACCAACCAACCACTGTTAAAGTTGTGACAGGTGGAAAAGCTTTCA 366
QY 1430 TTTAAGTGTATGTCATGTGACAGGACCAAGCCTTGGCGGAGCTTCTTTGTCGCTC 1489
DB 367 TTTAAGTGTATGTCATGTGACAGGACCAAGCCTTGGCGGAGCTTCTTTGTCGCTC 426
QY 1490 GTGATCAAAACGACGACACGCA 1512
DB 427 GTGATCAAAACGACGACACGCA 449

```

COMMENT

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
CONTACT: Weinstock, GM
Department of Biochemistry and Molecular Biology
University of Texas Medical School
6431 Fannin Street, Houston, TX 77030, USA
Tel: 713 500 6083
Fax: 713 500 0652
Email: georgiew@utmsg.med.utah.edu
Seq primer: pBluescript SK(-) c3

RESULT 6

B07889 400 bp DNA linear GSS 13-MAR-1998
LOCUS YX98T3 pBluescript expression library of genomic DNA from OG1RF
DEFINITION Enterococcus faecalis genomic clone YX98, genomic survey sequence.
ACCESSION B07889
VERSION B07889.1 GI:2058779

KEYWORDS GSS.
SOURCE Enterococcus faecalis
ORGANISM Enterococcus faecalis
Bacteria; Firmicutes; Lactobacillales; Enterococcaceae;
Enterococcus.
REFERENCE 1 (bases 1 to 400)
AUTHORS Xu, Y., Jiang, L., Murray, B.E. and Weinstein, G.M.
TITLE Enterococcus faecalis antigens in human infections
JOURNAL Infect. Immun. 65 (10), 4207-4215 (1997)
PUBMED 9317028
COMMENT Contact: Weinstein, GM
Department of Biochemistry and Molecular Biology
University of Texas Medical School
6431 Fannin Street, Houston, TX 77030, USA
Tel: 713 500 6083
Fax: 713 500 0652
Email: georgew@utmsi.med.utah.tmc.edu
Seq primer: pbluescript SK (-) c3
Class: shotgun
High quality sequence stop: 400.
Location/Qualifiers
1..400
/organism="Enterococcus faecalis"
/mol_type="genomic DNA"
/strain="OG1RF"
/db_xref="taxon:1351"
/clone="YX98"
/lab_host="XL1Blue MRF"
/clone_1fb="pbluescript expression library of genomic DNA
from OG1RF"
/note="Vector: pBluescript SK(-); site 1: HindIII; site 2:
EcoRI; Genomic DNA from an Enterococcus faecalis strain
OG1RF was partially digested with Sau3AI, sized
fractionated and ligated into the BamHI site of cosmid
vectors pLARX and pBel0AC11. After in vitro packaging
and plating, the cosmid clones were screened with antisera
from 4 enterococcal endocarditis patients and a rabbit
immunized with surface proteins from an E. faecalis
isolate. Cosmid DNA from immunopositive cosmid clones was
treated with Dnae I, the ends were filled with the Klenow
fragment, and then ligated into the EcoRV site in
pBluescript SK(-). The ligation mixture was transformed
into XL1Blue MRF, and the transformants were screened
with antisera. The immunopositive clones were selected for
DNA sequencing."

ORIGIN

Query Match 20.2%; Score 380; DB 11; Length 400;
Best Local Similarity 98.3%; Pred. No. 1.7e-78;
Matches 394; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 1032 TCMAATTTCTGTAATATTCCTCATTTGGGATTGCAACAAGAGCGCTATATAA 1091
DB 1 TCMAATTTCTGTAATATTCCTCATTTGGGATTGCAACAAGAA-GCGACGCTAATAA 59
QY 1092 CGTCAATTCATTTAGTGTGATTAACATATGACGCTTAACCTTTGATTAAGTGACTTC 1151
DB 60 CGTCAATTCATTTAGTGTGATTAACATATGACGCTTAACCTTTGATTAAGTGACTTC 119
QY 1152 TGGAGATGATGCTTATGCTTATATGATGGGATACAGTATGCTCTGAAAATTTATCA 1211
DB 120 TGGAGATGATGCTTATGCTTATATGATGGGATACAGTATGCTCTGAAAATTTATCA 179
QY 1212 AGTACTGAACAAGCAATAGCGCTTCACTGCGCGTTAATCCAGGCTAATTTCTACGCT 1271
DB 180 AGTACTGAACAAGCAATAGCGCTTCACTGCGCGTTAATCCAGGCTAATTTCTACGCT 239
QY 1272 AACACAGGCGGACACTAAATTTGTTACTTATGATTTAAATGAAGAAAGCAGATCC 1331
DB 240 AACGCGAGCGGACACTAAATTTGTTACTTATGATTTAAATGAAGAAAGCAGATCC 299
QY 1332 TACGAAGGCTTTAAATGAAGCGAATTTGATTAACGTCATCCGACGACCAACACC 1391
|||||

DB 300 TACGAAGGCTTTAAATGAAGCGAATTTGATTAACGTCATCCGACGACCAACACC 359
QY 1392 ACCAAGCTTTGAAGTTGACAGTGGGAAAGCTTTCAATTA 1432
DB 360 ACCAAGCTTTGAAGTTGACAGTGGGAAAGCTTTCAATTA 400
|||||

RESULT 7
A2548467
LOCUS A2548467 908 bp DNA linear GSS 14-NOV-2000
DEFINITION ENTERK30TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
genomic, genomic survey sequence.
ACCESSION A2548467
VERSION A2548467.1 GI:11172102
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 908)
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Reverse
Classes: Shotgun
High quality sequence start: 17
High quality sequence stop: 828.
Location/Qualifiers
1..908
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_1fb="Entamoeba histolytica Sheared DNA"
/note="Vector: PHOS1; Site 1: Bat I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + 1 method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. in Genome
Sequencing: A Practical Approach, eds M. Vaubin and B.
Barrell, Oxford University Press, 1999)."

FEATURES

source

Query Match 4.8%; Score 91.2; DB 11; Length 908;
Best Local Similarity 44.4%; Pred. No. 2.4e-10;
Matches 366; Conservative 0; Mismatches 458; Indels 0; Gaps 0;
QY 324 TGGGAACCTGTTGCTCAAGGAACCAACCGATGCAATGGGAATGACAGTTCACTTAC 383
DB 45 TGATGAAGAGATGATGAAGAAAGATATATGAAGATGATGATGAAGAA 104
QY 384 TAAAAAACAATATGTTAAAGATGACAGTGTATACATTAAGAAGAACCAAAAGAGGTGT 443
DB 105 TGATGAAGAGATGATGAAGAAAGATATATGAAGATGATGAAGAGATGA 164
QY 444 AGTTGCTGCTACGAATATAGTGTGCGCTCCAGCTTTACGAATGAATCAAGCAACACA 503
DB 165 AGACGACGAATATGAATTAAGAAAGATGATGATGAAGAAAGATGAATGAAGA 224
|||||

ORIGIN

QY	504		CGGTCCTTAATAATATGCAACGAAGATTAGCGGTGTCATATTTATCCTAAAAATGT	563
Db	225		TGATGAAGACCGATGGAAGACGACGAAATATGAAATTTAAGAGTATGATGATGAAGAAAGA	284
QY	564		GGTAGCCATGATGCTAGTTTACATGTGAAAAAATAGGAATCTGCTGAAAAATGAAGATT	623
Db	285		AGATGATGATGAAGAAGAAAGATGATGATGAAGATGATGAAGAAGAAAGATGAAGA	344
QY	624		AAATGCGCGAATTTGTTATTTCTPAAAAGCGAAGGCTCACAGGCAACGTAAAAATATAT	683
Db	345		AGATGATGAAGAAGATGATGAAGATGATGAAGAAAGAAAGATGATGATGAAGATGA	404
QY	684		CCAGAGATGCAAGATGATTTATATCATGCAACACGATTAAGAACCAAGCAAAACGCTT	743
Db	405		TGAAGACATGTAAGACGACGAATATGAATTTAGAAATGATGATGTAAGAAAGAAAGATGA	464
QY	744		TATTACTGGGAAAAAGTTATGAAATTTGCGCAAAATGATTTCAACAGAGCAGAGATGGAAC	803
Db	465		TGATGAAGAAGAAAGAAAGATGATGATGAAGATGATGATGAAGAAGAAAGATGATGA	524
QY	804		GGAGAAATTAACGTTAAAAATCTTGAAGTTGCTTCGTATTTTGAAGAAATGAAAAAC	863
Db	525		TGAAGAAAGATGTAAGATGATGATGAAGAAAGAAAGATGATGATGAAGAAAGATGAAGA	584
QY	864		TCCAATATATGAGAAATTAATGAAAAATCAAAACAAACACCATTTACAAATTGAACAA	923
Db	585		AGAGAAAGATGATGGAAGAAAGATGATGAAGAAAGAAAGATGATGATGAAGACGACGAATTA	644
QY	924		CAATCAAAACACCTGTTGAAAAAAACAGTCAAAATATATACCTTAAGTTGATPAAAAAC	983
Db	645		TGAATTAAGAAGTATGATGAAGAAGATGATGAAGAATGATGAAGAAGAAAGATGAAGA	704
QY	984		ACCAAGCTTAGATGCTAAGATGTGCGCAATTCGCGAAAAATTAATATCAAATTTCTGT	1043
Db	705		AGAAAGATGATGATGATGATGATGAAGAAAGATGATGATGAAGATGATGAAGAAAGATGA	764
QY	1044		AAATTTCCATTTGGGATTTGCGACAAAGAAAGCGACGCTAATAATCGTCAATTCA	1103
Db	765		AGATGATGAAGTATGATGAAGAAGATGAAGATGATGAAGATGATGAAGATGATGAAGATGA	824
QY	1104		TTTAGTTGATAAACAATGATGACGCTTAACTTTTGATTAACGTGA	1147
Db	825		TGAAGATGATGAAGACGACGAATTTGAATTTAAGATGATGA	868
RESULT 8				
LOCUS	AZ551618/c	843 bp	DNA	linear
DEFINITION	ENTDP54TR Entamoeba histolytica Sheared DNA			
ACCESSION	AZ551618			
VERSION	AZ551618.1			
KEYWORDS	GSS.			
SOURCE	Entamoeba histolytica			
ORGANISM	Entamoeba histolytica			
REFERENCE	1 (bases 1 to 843)			
AUTHORS	Loftus,B., Van Aken,S. and Fraser,C.			
TITLE	Determination of clone end sequences from Entamoeba histolytica			
JOURNAL	HMI:IMSS sheared DNA library			
COMMENT	Unpublished (2000)			
	Contact: Brendan J Loftus			
	Department of Eukaryotic Genomics			
	The Institute for Genomic Research			
	9712 Medical Center Dr., Rockville, MD 20850, USA			
	Tel: 301 838 0208			
	Fax: 301 838 3543			
	Email: bjl@fuser@igf.org			
	Clones are derived from the Entamoeba histolytica HMI:IMSS sheared			
	DNA library			
	Seq primer: ML3-Reverse			
	Class: shotgun			

```

High quality sequence start: 39
High quality sequence stop: 838
Location/Qualifiers
1. .843
FEATURES
source
```

```

/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/submitter="HMI:IMMS"
/db_xref="taxon:5759"
/clone_id="Entamoeba histolytica Sheared DNA"
/vector="pHOS1; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barelli, Oxford University Press, 1993). "
```

ORIGIN

Query Match	4.7%	Score 88.8	DB 11	Length 843
Best Local Similarly	44.5%	Pred. No. 8.7e-10		
Matches 351; Conservative	0;	Mismatches 437;	Indels 0;	Gaps 0

Qy	335	TTTGCTCAAGAACCA	CCGATGCAAAATGG	GAAATGTCAC	TGTCCTTACCTTA	AAAAACAA	334	
Db	822	TTGATGTATGAAGA	AGATGATGATGAAGA	AGATATGATGTATGA	GAAGAAGATGATGAAGA	G	763	
Qy	395	ATGCTAAGATGAC	AGTGTATCACTTAA	AGAAACCAAAAGG	GGGTGTGTGCTTA	454		
Db	762	ATGATGAAGAAGA	AGGTATGATGAAGA	TGATGAAGACGATGA	AGACGAGAT	703		
Qy	455	CGAATATGCTGTG	CGCTTCCCACTTAC	CAAAATGATCA	AGCAACAGATGTTCTATTA	514		
Db	702	ATGAAATTAAGA	AGATGATGATGAAGA	AGAAAGATGTATGA	GAAGATGAAGA	CG	643	
Qy	515	AATATGAACAG	AGAAATTTACGGT	TGCTTATTTATCCTTA	AAAAATGTGTAGCCAA	TG	574	
Db	642	ATGAAGACGACG	AAATATGAATTTGA	AGATATGATGATGA	GAAGAAGATGATGATG	583		
Qy	575	ATGCTAATTTAC	ATGTGA AAAAGT	AGAACTCTGTA	AAAAATGAAGATTAAT	TGGCGACG	634	
Db	582	AAGAAGCAAGA	AGATGATGAAGA	TGATGATGAAGA	AGAAAGATGATGAAG	523		
Qy	635	AATTTGTTATTT	CTAAAGCGAAGCG	CTCACAGCGACAGT	AAATATATCCAA	CGAGTCA	634	
Db	522	AAGATGTATGA	AGATGATGAAGA	AGAAAGATGATGTATGA	AGATGATGAAGA	CCATG	463	
Qy	695	AAGTGTATTA	TATACATGACAA	CGGATTAAGAACCA	AGCAAAACGCTTATTA	ATCTGGA	754	
Db	462	AAGACGACG	AAATATGAATTTGA	AGATGATGATGAAGA	AGAAAGATGATGAAGA	G	403	
Qy	755	AAAGTTATGA	AATTTGCGGAAAT	GAATTTACAGAGCAG	AGATGGAACGGGAA	ATTAA	814	
Db	402	AAGAAGATGTAT	GAAGAATGATGAAGA	AGAAAGATGATGTATGA	GAAGAAGA	G	343	
Qy	815	CAGTTAA	AAATCTTGAAGTTG	TGCTTGTATTTTAGA	AGAGTA	AAAGCTCCAAATTA	TAG	874
Db	342	ATGAAGTATGAT	GAAGAAAGAAAGATGAT	ATGAAGAGAAAGTGA	GAAGAAGAAAGATG	283		
Qy	875	CAGATTAAT	TGAAAATCAAA	CAAAAACACCATTT	CAATTGAAGCAAA	CAATCAAA	CAC	934
Db	282	ATGAAGAAGAT	GTATGAAGAAGAAAGATGAT	GAAGCGCAATAT	TGAATTTGAAG	223		
Qy	935	CTGTGAAAAA	ACAATCAAAAATG	ATATCTCTTAAGTT	ATATAACAAC	CAAGCTTG	934	
Db	222	ATGATGAAGA	AGATGATGAAGA	AGATGATGAAGA	AGATGTATGAAGA	GAAGAAGATG	163	
Qy	995	ATGCTAAGAT	TGCGCAATTTGCGCAAA	AAATTAATCAATTTCTGTA	AAATTTCCAT	1054		

Db 162 ATGATGATGATGATGAAGATGATGATGAAGATGATGAAGATGATGAAG 103
Qy 1055 TGGGATTCGACACAAAGAGCGACGCTAATTAATCTCAATTCATTTAGTTGATA 1114
Db 102 AAGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAG 43
Qy 1115 AACATGAT 1122
Db 42 AATATTAAT 35

RESULT 9
CNS06PLB/C
LOCUS T7 end of clone AV0AA015D02 of library AV0AA from strain CBS 379 of
DEFINITION Saccharomyces exiguus, genomic survey sequence.
ACCESSION AL409509
VERSION AL409509.1 GI:12177058
KEYWORDS GSS.
SOURCE Saccharomyces exiguus
ORGANISM Saccharomyces exiguus
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
AUTHORS Soulier,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Boulotin-Pukhara,M., Bon,B., Brotier,P., Casaregola,S.,
de-Montigny,D., Dujon,B., Durren,P., Lepingle,A., Liorente,B.,
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toffano-Nicoche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissbach,J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
PUBMED 11152876
REFERENCE 2 (bases 1 to 1084)
AUTHORS Bon,B., Neuvéglise,C., Lepingle,A., Wincker,P., Artiguenave,F.,
Galliardin,C. and Casaregola,S.
TITLE Genomic exploration of the hemiascomycetous yeasts: 6.
Saccharomyces exiguus
JOURNAL FEBS Lett. 487 (1), 42-46 (2000)
PUBMED 11152881
REFERENCE 3 (bases 1 to 1084)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbolicophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

FEATURES
source
1..1084
/organism="Saccharomyces exiguus"
/mol_type="genomic DNA"
/strain="CBS 379"
/db_xref="taxon:34358"
/clone="AV0AA015D02"
/clone_1b="AV0AA"
/note="end : 77"
misc_feature (<17..>979)
/inference="non-experimental evidence, no additional
details recorded"
/note="similar to Saccharomyces cerevisiae ORF YMR304w" [
UBP15 ; similarity to human ubiquitin-specific protease]
1 putative frameshift(s)"

ORIGIN
Query Match 4.2%; Score 79; DB 14; Length 1084;
Best Local Similarity 45.2%; Pred. No. 1.9e-07;
Matches 214; Conservative 17; Mismatches 242; Indels 0; Gaps 0;

Qy 670 ACAATTAATATATATCAAGAGTCAAGATGATTAATATCAATGCAACGATTAAGAA 729
Db 556 AA 497
Qy 730 CAAGCAAAACGCTTATTAATCTGGGAAAGATTATGAAATTCGGGAAATGATTCACAGAA 789
Db 436 AA 437
Qy 790 CGAGGAATGCAAGCGGGAATTAACATTAATAAATCTTGAGTTGTTGATATTTTA 849
Db 436 AA 377
Qy 850 GAAGAAATTAAGCTCCAAATTAATGCAATTAATGAAATCAAAACAAACACCATTT 909
Db 376 GAAA 317
Qy 910 ACAATTGAAGCAACATCAAAACACCTGTTGAAAAAACAATGAAATGATCTCTTAA 969
Db 316 AA 257
Qy 970 GTTGAATTAACACCAAGCTTAAGTGAAGATGTCGAATTCGCGAAAAATTA 1029
Db 256 AA 197
Qy 1030 TATCAATTTCTGTAATATATTCATTTGGGATTCGACAAAGAGCGACCTATTA 1089
Db 196 AA 137
Qy 1090 TACGTCAATTCATTTAGTTGATTAACATGATGACGCTTACTTTGATTA 1142
Db 136 GARGAAGATTAATGATTAATAATGTTCTATTCGAAATTAATTAATTTGATTA 84

RESULT 10
CG753732 1811 bp DNA linear GSS 24-OCT-2003
LOCUS CG753732/C
DEFINITION P048-4-G03.ya Ppa EcORI BAC library Pristionchus pacificus genomic,
genomic survey sequence.
ACCESSION CG753732
VERSION CG753732.1 GI:37978509
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasterida; Pristionchus.
REFERENCE 1 (bases 1 to 1811)
AUTHORS Srinivasan,J., Sins,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
PUBMED 12884007
COMMENT Contact: Sommer R.J
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES
source
1..1811
Location/Qualifiers
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_1b="Ppa EcORI BAC library"

ORIGIN

/note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."

Query Match 4.2%; Score 78.2; DB 12; Length 1811;
 Best Local Similarity 37.8%; Pred. No. 3.2e-07;
 Matches 248; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

OY 381 ACTTAAACAAATGCTAAAGATGCGATGATCCATTAAAGAAACCAAGAGGG 440
 DB 1712 ANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1653
 OY 441 TGTAGTCTGCTAGCAATATGCTGTCGCTCCCATGTTACGAATGATCAACCAAC 500
 DB 1652 ANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1593
 OY 501 AGATGCTCTTATATATGCAACAGAAATTTAGCGGTTGTCATATTTATCTTAAAA 560
 DB 1592 AA 1533
 OY 561 TGTGTAGCCATGATGTGTTTACATGTGAAAAAGTAGAAGCTCTGAAATGAAAG 620
 DB 1532 AA 1473
 OY 621 ATTAAATGCGCAGAAATTTGTTATTTCTTAAAGCAAGCTCACAGGCAAGTAAAT 680
 DB 1472 ANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1413
 OY 681 TATCCAGAGTCAAGATGATGATATATATCATGACACGATTAAGAAACGAAAAAG 740
 DB 1412 AA 1353
 OY 741 CTTATTACTGGGAAAGTTATGAAATGCGCAAAATGATTTCAAGAGCAGAGATGG 800
 DB 1352 NNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1293
 OY 801 AACCGAGATTAACAGTTAAATCTTGAGGTGTTGTTATTTTGAAGAAAGTAA 860
 DB 1292 AA 1233
 OY 861 AGCTCAATATATGAGATTTATGAAATCAACAAACACCATTTTACATTTGAAGC 920
 DB 1232 NNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNAN 1173
 OY 921 AAACATCAACACCTGTTGAAAAACAGTCAAAATGATCTCTTAAAGTTGATTAAC 980
 DB 1172 AA 1113
 OY 981 AACACCAAGCTTGAATGTAAGATGTGCAATTTGGCAAAATTAATATCAAA 1036
 DB 1112 AA 1057

RESULT 11
 LOCUS AZ531291 877 bp DNA linear GSS 03-NOV-2000
 DEFINITION ENTBQ34TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
 genome, genomic survey sequence.

ACCESSION AZ531291
 VERSION AZ531291.1 GI:11085838
 KEYWORDS GSS.

SOURCE Entamoeba histolytica

ORGANISM Entamoeba histolytica
 Entamoeba histolytica
 Eukaryota; Entamoebidae; Entamoeba.
 1 (bases 1 to 877)
 Reference: Loftus, B., Van Aken, S., and Fraser, C.
 Determination of clone end sequences from Entamoeba histolytica
 title HM:IMSS sheared DNA library

JOURNAL COMMENT
 Unpublished (2000)
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research

FEATURES

source
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjoftus@ciqr.org
 Clones are derived from the Entamoeba histolytica HM:IMSS sheared
 DNA library
 Seq primer: MJ3-Reverse
 Class: shotgun
 High quality sequence start: 22
 High quality sequence stop: 829.
 location/Qualifiers
 1..877
 /organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HM:IMSS"
 /db_xref="taxon:5759"
 /clone_id="Entamoeba histolytica Sheared DNA"
 /note="vector: pHO51; Site 1: Bst I; Constructed at the
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica
 using a method described by Clark and Diamond (Clark,
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
 method for isolate identification. Exp. Parasitol.
 77:450.). The DNA was mechanically sheared to give a
 tight size distribution (~2 kb). The v + i method used for
 the library construction is described in detail in Smith,
 H.O. and Venter, J.C. (Making small insert libraries for
 whole genome shotgun sequencing projects. in Genome
 Sequencing: A Practical Approach, ed. M. Vaubin and B.
 Barrell, Oxford University Press, 1999)."

ORIGIN

Query Match 4.1%; Score 78; DB 11; Length 877;
 Best Local Similarity 45.2%; Pred. No. 3.1e-07;
 Matches 328; Conservative 0; Mismatches 395; Indels 3; Gaps 1;

OY 251 ACGAATTTTACGAGCAAGCGGCGGCAAGCGTGTGACGCTTAAACAGCTCTCC 310
 DB 805 ACGATATGATGATGAGAAAGAGATGATGATGATGATGATGATGATGATGATG 746
 OY 311 AAAGTTAACTCTGGGAAACCTGTTCTCAAGAAACCAAGATCAAAATGGAAATGCA 370
 DB 745 AAGATGATGATGATGAGAAAGATGATGATGATGATGATGATGATGATGATG 686
 OY 371 CTGTTCACTTACCTTAAACAAATGCTAAAGATGATGATGATGATGATGATGATG 430
 DB 685 ATGAAGAAAGATGATGAGAAAGATGATGAGAAAGAAAGATGATGATGATGATG 626
 OY 431 CAAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 490
 DB 625 AAGACATGAAGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 566
 OY 491 TCAAGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 550
 DB 565 ATGATGAAGATGATGAGAAAGATGATGAGAAAGATGATGATGATGATGATGATG 506
 OY 551 ATCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 610
 DB 505 ATGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 449
 OY 611 AAAATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 670
 DB 448 AAGAGATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 389
 OY 671 CAGTAAATATATCCAGAGATCAAAAGATGATGATGATGATGATGATGATGATG 730
 DB 388 ATGATGAAGATGATGAGAAAGATGATGAGAAAGATGATGATGATGATGATGATG 329
 OY 731 AAGCAAAACGCTTTATCTGGGAAAGATGATGATGATGATGATGATGATGATG 790
 DB 328 AAGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 269
 OY 791 CAGAGAAATGAACGAGAAATTAACATTAATCTTGAAGTGTGCTGATATTTTAA 850

Db 268 ATGAAGATGATGATGATGAAGAGAGATGATGATGAAGAGAGATGATGAAG 209
Qy 851 AAGAGTAAGAGCTCCAAATTAATGCAATTAATGAAAAATCAACAAACACATTTA 910
Db 208 AAGAAGATGAAGAGAGAGATGATGAAGAGATGATGAAGAGAGATGATGAAG 149
Qy 911 CAATTGAAGCAACATCAACACCTGTTGAAAAAAGTCAAAATGATTAAG 970
Db 148 AAGAGCAGCAATTAATTAATTAAGATGATGAAGAGATGATGAAGAGATGAAG 89
Qy 971 TTGATA 976
Db 88 AAGATA 83

RESULT 12
DN711652/c 1416 bp mRNA linear EST 30-MAR-2005
LOCUS CLJ96-E12 5', mRNA sequence.
DEFINITION DN711652
ACCESSION DN711652
VERSION DN711652.1 GI:62076691
KEYWORDS EST.
SOURCE Gasterosteus aculeatus (three spined stickleback).
ORGANISM Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 1416)
Kingsley, P.M., Petchel, C., Balabhadra, S., Grimwood, J., Dickson, M.,
Schmutz, J., and Myers, R.M.
Expressed sequence tags from Gasterosteus aculeatus
Unpublished (2003)
Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@hgsc.stanford.edu
Plate: 96
High quality sequence start: 30
High quality sequence stop: 303.
Location/Qualifiers
1..1416
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Bilrnfjordur marine sticklebacks, Iceland"
/db_xref="taxon:69293"
/clone="CLJ96-E12"
/sex="mixed male and female"
/tissue_type="whole larva"
/dev_stage="21 day old larvae collected at Swarup Stage 30
(U. Embryol. Exp. Morphol 6: 373-383, 1958)"
/lab_host="DH10B (T1 phage resistant)"
/clone_11b="SHGC-CLJ"
/note="Vector: Express 1; Total and poly A+ RNA was
isolated from the indicated stickleback tissue, and a cDNA
library was constructed in the Express 1 plasmid vector by
Open Biosystems. First strand cDNA synthesis was primed
with an 54 bp linker primer containing an oligodT sequence
preceded by a synthetic NotI site (first strand primer:
5'-GACTAGTCTAGATCGGAGCGCGCCGCTT-3'). Following
second strand synthesis, cDNAs were made blunt at the end
corresponding to the original 5' prime end of mRNA, and
cloned directionally into the NotI and EcoRV sites of
Express 1. Note that the EcoRV site is typically destroyed
in the blunt end cloning, leaving a junction of the form
'xxATC' (where is ATC is the second half of the EcoRV
site, and xxx is derived from the cDNA sequence). A map of
the Express 1 vector is available at:

ORIGIN

http://www.openbiosystems.com/cdna_library_construction_fa
g.php#8 The primary library was transformed and amplified
in DH10B (T1 phage resistant) Bacteria. Clones available
from Open Biosystems:
http://www.openbiosystems.com/stickleback"

Query Match 4.1%; Score 77.6; DB 9; Length 1416;
Best Local Similarity 39.1%; Pred. No. 4.2e-07;
Matches 248; Conservative 0; Mismatches 386; Indels 0; Gaps 0;

Qy 485 AATGATCAAGCAACAGATGGTCCCTTAATTAATGACAGAGATAGCGGTGTC 544
Db 1122 AA 1063
Qy 545 ATATTATCTTAATAATGTGTAGCCATGATGTTTCACTGATGAAGAGTGA 604
Db 1062 NNN 1003
Qy 605 CTGCTGAATAATGAGATTAATGCGCGCAATTTGTTATTTCTAAGCGAAGCTC 664
Db 1002 NNN 943
Qy 665 CAGCAGCAATTAATTAATCAAGAGTCAAGATGATTAATTAATCAAGCAGATA 724
Db 942 AA 883
Qy 725 AAGAACAGCAACACGCTTATTACTGGAAGATTGAATTTGGCGAATGATTCA 784
Db 882 AA 823
Qy 785 CAGAACAGCAATGGAACGGAGATTAACTGTAATTTGAGTTGTTGCTATA 844
Db 822 NNN 763
Qy 845 TTTTGAAGAGTAAAGCTCCAAATTAATGCAATTAATGAAATCAACAAAAAC 904
Db 762 AA 703
Qy 905 CATTTCAATTGAGCAACCAATCAACACCTGTTGAAAAACAGTCAAAATGATACCT 964
Db 702 AA 643
Qy 965 CTAAAGTTGATTAACACACCACTTGAATGATGATGATGATGATGATGAT 1024
Db 642 AA 583
Qy 1025 TTAATATCAATTTCTGTAAATTTCCATTGGGATTCAGACAAAGAGCGACTA 1084
Db 582 AA 523
Qy 1085 ATAAATACCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1118
Db 522 AA 489

RESULT 13
AZ551092 912 bp DNA linear GSS 14-NOV-2000
LOCUS ENTJ22TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, genomic survey sequence.
ACCESSION AZ551092
VERSION AZ551092.1 GI:11176393
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 912)
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.
TITLE Determination of clone end sequences from Entamoeba histolytica
JOURNAL HMI:IMSS sheared DNA library
COMMENT Unpublished (2000)
Contact: Brendan J Loftus

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjoftus@igr.org

Clones are derived from the Entamoeba histolytica HMI-IMSS sheared DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 861.
Location/Qualifiers

FEATURES

source

```
1..912
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHOSt; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."
```

ORIGIN

```
Query Match      4.1%; Score 77.2; DB 11; Length 912;
Best Local Similarity 44.3%; Pred. No. 4.8e-07;
Matches 316; Conservative 0; Mismatches 398; Indels 0; Gaps 0;

QY 518 ATGAAACGAGAAATTTAGCGGTTGTCATTTTCTTAAATGTGTCGCAATGATG 577
DB 136 AAGAAGAAATGAAATTAAGAAAACAATCAACAAATTAAGAAAGAAAGTGAAGCG 195
QY 578 GTAGTTTACATGTGAAAAAGTGAAGAACTGTAATTAAGATTAATGCGCGAAT 637
DB 196 AATATGATTTAGAAAGAAAGAAAGATGATGATGATGATGAAAGAAAGATGATG 255
QY 638 TTGTTATTTCTAAAAGCGAAGGCTCACGACAGCAATTAATATATCCAAAGTCAAG 697
DB 256 ATGATGATGATGAAAGAAAGATGATGATGATGATGATGATGATGATGATG 315
QY 698 ATGATTTATTCATGACACGATTAAGAAACAAGCAAAAGCTTTATTTACTGGAAA 757
DB 316 AAGAAGATGATGAAAGAAAGTGAAGAAAGAAAGATGATGATGATGATGATG 375
QY 758 GTTATGAATTTGCGCAAAATGATTTCAACGAGACAGAGATGAACGGAATTAACG 817
DB 376 ACGATGAAGACGACCAATTTGATTAAGAAAGATGATGATGATGATGATGATG 435
QY 818 TTTAAATCTTGAGGTTGTTGTTATTTTGAAGAAAGTAAAGCTCCAAATTAAGCG 877
DB 436 ATGAAGATGATGAAAGCAAGTGAAGACGATGATGATGATGATGATGATGATG 495
QY 878 AATTATTTGAAATCAAAACAAAACCCCTTTTACATTTGAAGCAAAACATCAACACCTG 937
DB 496 AAGAAGAAAGATGATGAAAGAAAGAAAGATGATGATGATGATGATGATG 555
QY 938 TTTAAAAACAGTCAAAATGATACCTTAAAGTTGATTAACCAACACCAAGCTTAGATG 997
DB 556 ATGAAGAAAGATGATGAAAGAAAGTGAAGATGATGATGATGATGATGATGATG 615
QY 998 GTAAAGATGTGGCAATTTGGCGAAAAATTTAAATATCAATTTCTGTAATATTTCCATTGG 1057
DB 616 AAGATGATGAAGACATGAAAGACGACGATATGAATTTGAAGATGATGATGATG 675
```

```
QY 1058 GGATTCAGACAAAGAGCGACGCTAATTAATAGTCAATTTAGTTGATTAAC 1117
DB 676 AAGATGAAGATGAAGAAAGAAAGATGATGATGATGATGATGATGATGATG 735
QY 1118 ATGATGACAGCCCTTAATCTTTGATTAACGTAAGTCTTGAGAGATGATCTTAATGCTTATATG 1177
DB 736 ATGATGATGAAGAAAGAAAGATGATGATGATGATGATGATGATGATGATG 795
QY 1178 ATGGGATACAGTATTTGCTCTGAAAATTTATCAAGTGAATGACAAAGCAATG 1231
DB 796 ATGAAGAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 849
```

RESULT 14

```
CNS01JRG/c 879 bp DNA linear GSS 12-JUN-2001
LOCUS
DEFINITION
```

Anopheles gambiae GSS 17 end of clone 14D07 of NotreDame1 library from strain PST of Anopheles gambiae (African malaria mosquito), genomic survey sequence.

```
ALI47405
ALI47405.1 GI:7005551
```

```
VERSION
KEYWORDS
SOURCE
ORGANISM
```

Anopheles gambiae (African malaria mosquito)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae; Anophelinae; Anopheles.

```
REFERENCE
AUTHORS
TITLE
```

Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)

```
JOURNAL
REFERENCE
AUTHORS
TITLE
```

Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)

```
JOURNAL
REFERENCE
AUTHORS
TITLE
```

Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)

```
JOURNAL
REFERENCE
AUTHORS
TITLE
```

Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)

```
JOURNAL
REFERENCE
AUTHORS
TITLE
```

Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)

```
JOURNAL
REFERENCE
AUTHORS
TITLE
```

Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)

```
JOURNAL
REFERENCE
AUTHORS
TITLE
```

Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)

```
JOURNAL
REFERENCE
AUTHORS
TITLE
```

Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)

```
JOURNAL
REFERENCE
AUTHORS
TITLE
```

Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)

```
JOURNAL
REFERENCE
AUTHORS
TITLE
```

Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)

Qy	623	TAAATGGCGCAATTGTTATTTCTAAAGGAAGCGCACAGCAAGAAATATA	682
Db	552	AAKAAAAAAAAAAKGTTRARAAATAAAAAAAAAKKTRRAAAAAAAAAAKGDAAR	683
Qy	683	TCCAAGAGTCMAAGATGATTTATATACATGCAACAGGATTAAGAACACGCAAA	742
Db	492	WAGAAAAAAAAAGATWYKAKKAAAAAAAAAAAAAAAAATDAAAAAAAAATYAMTWBAAAAA	743
Qy	743	TTATTTACTGGGAAAAAGTTATGAAATTTGGCGAAATGATTTCCACAGACAGAA	802
Db	432	ATAKTTGTRAAAAAAAAADAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	803
Qy	803	CGGAGAAATTAACTGTTAAAACTTGAGTTGTTGCTATATTTAGAGAGTAAAG	862
Db	372	AA	863
Qy	863	CTCCAAATPAATGCAGAATTAAATTGAAATCAACAAACACCAATTCAATTGAAGA	922
Db	312	AA	923
Qy	923	ACAAATCAACAACCTGTTGAAAAAACGTCAAAATGATACCTTAAGTTGATAAACA	982
Db	252	AA	983
Qy	983	CACCAAGTTAGATGTAAGATGTGGCAATTGGCGAAAAATTAAATTCAAATTTCTG	1042
Db	192	AA	1043
Qy	1043	TAAATATTCATTGGGAGTTGCAGACAAGAAAGCGCAAGCTAATTAATACGTCAATTCA	1102
Db	132	AAATTAATAAAAAAAAAATATAAAAAAAAAAAAAAAAANNNNNNNNNAANNCABMTTTANANA	1103
Qy	1103	ATTTAGTTGATTAACA	1118
Db	72	ANNNTATNTAAAAA	57

Query Match	4.1%	Score 76.4	DB 7	Length 993
Best Local Similarity	49.2%	Pred. No. 7.5e-07		
Matches 246	Conservative 0	Mismatches 252	Indels 2	Gaps 2
QY	645	TTCTAAAAAGCGAAGCTCAACGAGCAGATATAATATATCCAGAGTCAAAAGATG	704	
DB	111	TTGCGCCAGCAAAAGACTCTTTAGGTTCAAAAAGAAAACCTGCGAGATATTATGTTGAA	170	
QY	705	ATATACATGACAAACGATTAAGAACAAAGCAAAACGCTTTATTTCTGGGAAAATGTA	764	
DB	171	ACTNAAAAAAAAAAAAAAAAAAAAAAAAATTTAAATTTTTTTA-TTATTAANNNTTAA	229	
QY	765	AATTGGGAAATGATTTTCAAGAGCAGAAATGGAACGGAGAAATTAACGTTAAAA	824	
DB	230	AAAAAAAAATTTTTAA	285	
QY	825	TCTTGAGTGGTTCGTATATTTTACAGAGATTAAGCTCCAATATATGCGAATTAAT	884	
DB	290	AAAAAAAAATTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAA	349	
QY	885	TGAAAATCAAAACAAAAACACCTTTACAATTGAAGCAACAAATGAAACCTGTTGAAA	944	
DB	350	AAAAAAAAAAAAAAAAAAAAAGATTAATTAATAAAAAAAAAA-AAAAAAAAAAAAACCTTTAATAT	408	
QY	945	AACAGTAAAAATATATCTCTAAGTTGATTAAGTAAACAACCAAGCTTAGTGTAAGA	1004	
DB	409	AATTAATAAAAAAAAAATACGCAAAAAAAAAAAAAAAAAAAAAATATTAATAAAAAAAAAA	468	
QY	1005	TGTGCAATGGCGAATAAAATTAATAATATCAAAATTTCTGTAATATATTCATTTGGGATTC	1064	
DB	469	AAAAGAAATCCGCAAAATATTTATATATTTTAAATTTTAAATATTAACGAAAAAATA	528	
QY	1065	AGACAAAGAAAGCGACGCTATATAATACGTCAATTCATTTAGTATTAACATGATGC	1124	
DB	529	AAAAAAAAAAAAAAAAATAATATATATATGATTAATAATTAATAATATATATATATTA	588	
QY	1125	AGCCTTAACCTTTGATTAACG 1144		
DB	589	AAAAAAAAAATTTGAAAAGG 608		

Search completed: July 5, 2006, 22:06:38
Job time : 6170 secs

FEATURES	Location/Qualifiers
source	1. .993

ORIGIN

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 18:05:13 ; Search time 826 Seconds
(without alignments)
15877.481 Million cell updates/sec

Title: US-10-661-809a-12

Perfect score: 1881

Sequence: 1 atgaagcaatcaaaaaagc.....gacgragaagaagaatgct 1881

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_8:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*
15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1881	100.0	1980	10	ADH84834
2	1881	100.0	1983	2	AAK20108
3	1881	100.0	1983	6	ABN98093
4	1881	100.0	1983	6	ACA88057
5	1881	100.0	1983	8	ABX61663
6	1881	100.0	1983	14	ADY38951
7	1881	100.0	15614	2	AAK12982
8	1881	100.0	15614	6	ABN98777
9	1859	98.8	1875	12	ADH84888
10	1859	98.8	1875	14	ADV16333
11	1687	89.7	1687	2	AAK20109
12	1687	89.7	1687	6	ABN98094
13	1687	89.7	1687	8	ACA88058
14	1687	89.7	1687	8	ABX61664
15	1687	89.7	1687	8	ADY38953
16	1014.6	53.9	1878	12	ADH84894
17	1014.6	53.9	1878	14	ADV16652
18	1014.6	53.9	1995	10	ADC93971

19	595	31.6	2199	6	ABK11591	Abk11591 E. faecal
20	83.8	4.5	3399	2	AAO5868	Aao5868 Chicken 1
21	80	4.3	1701	14	ADZ7262	Adz7262 Plasmodu
22	79.8	4.2	2395	15	AE66299	Aee66299 Leucocy
23	70.4	3.7	1686	2	AAQ87587	Aaq87587 DNA encod
24	65	3.5	627	13	ACN54555	Acn54555 Cotton an
25	64.8	3.4	1998	8	AAA70212	Aaa70212 Plasmodu
26	64.6	3.4	6609	8	ABZ22899	Abz22899 Streptoc
27	64.2	3.4	2115	6	ABN67915	Abn67915 Streptoc
28	64.2	3.4	14006	6	ABJ33958	Abj33958 Human imm
29	64.2	3.4	11000	6	ABN71527_13	Abn71527_13 Contin
30	63.2	3.4	1509	5	AAO2390	Aao2390 Virulent
31	63.2	3.4	1509	5	ABK11585	Abk11585 S. agalac
32	63	3.3	822	6	ABO51819	AbO51819 Oligonuc
33	63	3.3	822	6	ABQ44852	AbQ44852 Oligonuc
34	63	3.3	822	6	ABQ44853	AbQ44853 Oligonuc
35	63	3.3	822	6	ABQ44853	AbQ44853 Oligonuc
36	63	3.3	3579	3	AAA70099	Aaa70099 Plasmodu
37	61.8	3.3	9539	4	AA645347	Aaa645347 Chemical
38	61.8	3.3	9539	6	ABK28180	Abk28180 DNA trans
39	61.6	3.3	516	8	ABX40620	Abx40620 Bovine ES
40	61	3.2	700	10	ACD92384	AcD92384 Human col
41	60.6	3.2	1065	14	ADZ72252	Adz72252 Plasmodu
42	60.6	3.2	1065	15	AEF80820	Aef80820 Plasmodu
43	60	3.2	15674	6	ABJ32363	Abj32363 Human imm
44	60	3.2	15674	6	ABJ34477	Abj34477 Human met
45	60	3.2	15674	6	ABJ70514	Abj70514 Chemical

ALIGNMENTS

RESULT 1	ADH84834	standard; DNA; 1980 BP.
ID	ADH84834	
XX	ADH84834;	
XX	22-APR-2004	(first entry)
XX	Enterococcus faecalis polynucleotide #2719.	
DE	Enterococcus faecalis infection; transcription regulatory element;	
XX	antibacterial; gene; de.	
KW	Enterococcus faecalis.	
XX	US6617156-B1.	
XX	09-SEP-2003.	
PD	13-AUG-1998;	98US-00134000.
XX	13-AUG-1998;	98US-00134000.
PF	15-AUG-1997;	97US-0055778P.
XX	15-AUG-1997;	97US-0055778P.
PR	(DOUC/) DOUCETTE-STAMM L A.	
PA	(BUSH/) BUSH D.	
XX	Doucette-Stamm LA, Bush D;	
PI	WPI; 2003-895934/82.	
XX	P-PSDB; ADH88239.	
DR	New nucleic acid comprising a sequence encoding an Enterococcus faecalis	
XX	polypeptide, useful for preparing a composition for diagnosing or	
PT	treating E. faecalis infection.	
XX	Dielosure; SEQ ID NO 2719; 193pp; English.	
PS	The invention relates to Enterococcus faecalis polynucleotides and	
CC	polypeptides. The invention also relates to a recombinant expression	
CC	vector comprising a polynucleotide operably linked to a transcription	
CC	regulatory element, a cell comprising a recombinant vector, a method for	

producing an E. faecalis polypeptide, an isolated nucleic acid comprising a sequence not given in the specification, a recombinant vector comprising the nucleic acid and a cell comprising the recombinant vector. The polynucleotides can be used to detect the presence of E. faecalis in a sample. The sequences are useful for preparing a composition for diagnosing or treating Enterococcus faecalis infection. This sequence represents an E. faecalis polynucleotide of the invention.

Sequence 1980 BP; 728 A; 330 C; 425 G; 497 T; 0 U; 0 Other;

Query Match 100.0%; Score 1881; DB 10; Length 1980;

Best Local Similarity 100.0%; Pred. No. 0; Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAAGCAATTAAAAAGTTTGGTACACCGTTAGTACTTGTGTACTAATTTTGGCAGCTT 60
DB 97 ATGAAGCAATTAAAAAGTTTGGTACACCGTTAGTACTTGTGTACTAATTTTGGCAGCTT 156
QY 61 TTCAAGGTATTTGGGCAACAACGTGATTTGCGAAGAAATGGGGAGCGCAG 120
DB 157 TTCAAGGTATTTGGGCAACAACGTGATTTGCGAAGAAATGGGGAGCGCAG 216
QY 121 CTGCGATTTCAAAAAAGAAAATGAACGATTTACAGATCCGCTTATTCAAAATAGCGGG 180
DB 217 CTGCGATTTCAAAAAAGAAAATGAACGATTTACAGATCCGCTTATTCAAAATAGCGGG 276
QY 181 AAGAAATGAGCGAGTTTGTATTAATATCAGGACTGGCAGATGTGACGTTTGTATTTAT 240
DB 277 AAGAAATGAGCGAGTTTGTATTAATATCAGGACTGGCAGATGTGACGTTTGTATTTAT 336
QY 241 AACGTGACGAACGAATTTTACAGACGAGCGGCGACGCAAGCGTTGATCAGCTTAA 300
DB 337 AACGTGACGAACGAATTTTACAGACGAGCGGCGACGCGCAAGCGTTGATCAGCTTAA 396
QY 301 CAAGCTGTCCAAAGTTTAACTCTCGGGAAAACCTGTTCTCAAGAAACCAACGATGCAAT 360
DB 397 CAAGCTGTCCAAAGTTTAACTCTCGGGAAAACCTGTTCTCAAGAAACCAACGATGCAAT 456
QY 361 GGGAAATGTCACTGTTCACTTACCTTAAAAACAAAATGTAAAGATGCACTGTATACAT 420
DB 457 GGGAAATGTCACTGTTCACTTACCTTAAAAACAAAATGTAAAGATGCACTGTATACAT 516
QY 421 AAGAAAGAACCAAAAAGAGGGTGTAGTGTGCTACGATATGATGCTGCTTCCAGTT 480
DB 517 AAGAAAGAACCAAAAAGAGGGTGTAGTGTGCTACGATATGATGCTGCTTCCAGTT 576
QY 481 TAGCAATGATCAGCAACAGATGGTCTCTATAATATATGAACAGAAAGATTAAGCGGTT 540
DB 577 TAGCAATGATCAGCAACAGATGGTCTCTATAATATATGAACAGAAAGATTAAGCGGTT 636
QY 541 GTTCAATTTATCTTAAAAATGTGGTAGCCATGATGTGATTTACATGTGAAAAAGTA 600
DB 637 GTTCAATTTATCTTAAAAATGTGGTAGCCATGATGTGATTTACATGTGAAAAAGTA 696
QY 601 GGAACGTGCAAAAATGAAGGATTTAAATGGCGAGAAATTTGTTTCTTAAAGCGAAGGC 660
DB 697 GGAACGTGCAAAAATGAAGGATTTAAATGGCGAGAAATTTGTTTCTTAAAGCGAAGGC 756
QY 661 TCACCAAGGACAGTAAATATATCCAAAGAGTCAAGATGATTAATATCATGCAACAG 720
DB 757 TCACCAAGGACAGTAAATATATCCAAAGAGTCAAGATGATTAATATCATGCAACAG 816
QY 721 GATTAAGAACAAAGCAAAACGCTTTATTAATCTGGGAAAAGTTATGAAATTTGGCGAAATGAT 780
DB 817 GATTAAGAACAAAGCAAAACGCTTTATTAATCTGGGAAAAGTTATGAAATTTGGCGAAATGAT 876
QY 781 TTCAAGAGAGCAGAAATGAAACGGGAGAAATTAACGTTAAAAATCTTGAAGTGGTTGG 840
DB 877 TTCAAGAGAGCAGAAATGAAACGGGAGAAATTAACGTTAAAAATCTTGAAGTGGTTGG 936
QY 841 TATATTTTGAAGAGTAAAGCTCCAAATTAATGCAAGATTAATGAAATCAAAACAAA 900
DB 937 TATATTTTGAAGAGTAAAGCTCCAAATTAATGCAAGATTAATGAAATCAAAACAAA 996
```

```
QY 901 ACACATTTACATTAAGAAAGCAACATCAAAACACCTGTGAAAAAAGTCAAAATGAT 960
DB 997 ACACATTTACATTAAGAAAGCAACATCAAAACACCTGTGAAAAAAGTCAAAATGAT 1056
QY 961 ACCTTAAGTTGATTAAGCAACACCAAGCTTGAATGATGTAAGATGTGGCAATTTGGCGAA 1020
DB 1057 ACCTTAAGTTGATTAAGCAACACCAAGCTTGAATGATGTAAGATGTGGCAATTTGGCGAA 1116
QY 1021 AAAATTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAAGAAAGCGGAC 1080
DB 1117 AAAATTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAAGAAAGCGGAC 1176
QY 1081 GCTAATTAATACGTCAATTCATTTAGTGAATAACATGATGACGCTTAACTTTGAT 1140
DB 1177 GCTAATTAATACGTCAATTCATTTAGTGAATAACATGATGACGCTTAACTTTGAT 1236
QY 1141 AACGTGACTTCTGAGAGATATGCTTATGCTTAAATATGATGGGATTAACATGATGCTCT 1200
DB 1237 AACGTGACTTCTGAGAGATATGCTTATGCTTAAATATGATGGGATTAACATGATGCTCT 1296
QY 1201 GAAATTTATCAAGTACTGAAACACCAAAATGAGCTTCACTGTGCGCTTAAATCCAGTAT 1260
DB 1297 GAAATTTATCAAGTACTGAAACACCAAAATGAGCTTCACTGTGCGCTTAAATCCAGTAT 1356
QY 1261 ATTCTAGGCTTAACACCAAGGCGGCACTAAATTCGTTTACTTATGCAATTAATGAA 1320
DB 1357 ATTCTAGGCTTAACACCAAGGCGGCACTAAATTCGTTTACTTATGCAATTAATGAA 1416
QY 1321 AAGCAGATCTTAACGAAGGCTTTAAAAATGAGCGAAATGTTGATACGGTCATACGAC 1380
DB 1417 AAGCAGATCTTAACGAAGGCTTTAAAAATGAGCGAAATGTTGATACGGTCATACGAC 1476
QY 1381 GACCAAAACCAACCACTGTTGAATTTGTGACAGTGGGAAAACGTTTCAATTAATGCAT 1440
DB 1477 GACCAAAACCAACCACTGTTGAATTTGTGACAGTGGGAAAACGTTTCAATTAATGCAT 1536
QY 1441 GGGCATGTGACGACGACCAAGCTTTGGCGGAGGCTTCTTTGTGCTGCTGATCAAAAC 1500
DB 1537 GGGCATGTGACGACGACCAAGCTTTGGCGGAGGCTTCTTTGTGCTGCTGATCAAAAC 1596
QY 1501 AGCGACACGCAAAATTAATTTGAATAATGCAAGAAACGAAAGCAGCACTTGGGTGAAA 1560
DB 1597 AGCGACACGCAAAATTAATTTGAATAATGCAAGAAACGAAAGCAGCACTTGGGTGAAA 1656
QY 1561 ACAAAGCTGAAGCACTTCTTTACAAACACGCTGATGATTTGATATCAACAGG 1620
DB 1657 ACAAAGCTGAAGCACTTCTTTACAAACACGCTGATGATTTGATATCAACAGG 1716
QY 1621 CTTAAATACGGTACCTAATTTAGAAAGAACTGTAGCTCTGATGATTAATGTCTTGT 1680
DB 1717 CTTAAATACGGTACCTAATTTAGAAAGAACTGTAGCTCTGATGATTAATGTCTTGT 1776
QY 1681 ACAAAATCGGATTAATTTGTGTGATGATGATCAATGATGCAACAGAAAACTTATG 1740
DB 1777 ACAAAATCGGATTAATTTGTGTGATGATGATCAATGATGCAACAGAAAACTTATG 1836
QY 1741 TCACCAAGAAAGTACCAACCAACCAAGTACTTACCTTCAACAGGTGGCAAGGA 1800
DB 1837 TCACCAAGAAAGTACCAACCAACCAAGTACTTACCTTCAACAGGTGGCAAGGA 1896
QY 1801 ATCTAGCTTACTTGAAGAGTGGCGAGCTGCTACTTATTAAGGAGAGTCACTTGTGCT 1860
DB 1897 ATCTAGCTTACTTGAAGAGTGGCGAGCTGCTACTTATTAAGGAGAGTCACTTGTGCT 1956
QY 1861 AGACGTAGAAAAAGAAATGCT 1881
DB 1957 AGACGTAGAAAAAGAAATGCT 1977
```

RESULT 2
AAK20108
ID AAK20108 standard; DNA; 1983 BP.

XX AC AAX20108;
 XX XX 20-APR-1999 (first entry)
 XX DE Enterococcus faecalis gene EF058.
 XX KM Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
 XX KM detection; attenuation; antigenic; ss.
 XX OS Enterococcus faecalis.
 XX PN M09850554-A2.
 XX XX 12-NOV-1998.
 XX PF 04-MAY-1998; 98WO-US008959.
 XX PR 06-MAY-1997; 97US-0044031P.
 XX PR 16-MAY-1997; 97US-004655P.
 XX PR 14-NOV-1997; 97US-0066009P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Kunsch CA, Choi GH, Bailey C, Hromockyj A;
 XX DR WPI; 1999-070095/06.
 XX DR P-PSDB; AAY00118.
 XX PT New isolated Enterococcus faecalis polynucleotides - used to develop
 PT products for the detection of Enterococcus and for use in vaccines for
 PT prevention or attenuation of Enterococcus infection.
 XX PS Claim 1; Page 133; 301pp; English.
 XX CC The present sequence represents a gene isolated from Enterococcus
 CC faecalis. The present invention describes genes, proteins and antigenic
 CC polypeptides isolated from B. faecalis. The proteins can be used in
 CC vaccines for preventing or attenuating an infection caused by a member of
 CC the Enterococcus genus in an animal. They can also be used for detecting
 CC Enterococcus antibodies in a sample. The nucleotide sequences can be used
 CC for detecting Enterococcus nucleic acids. Products from the present
 CC invention can also be used for screening compounds to identify agonists
 CC and antagonists of E. faecalis protein activity
 XX CC
 SQ Sequence 1983 BP; 729 A; 330 C; 426 G; 498 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1881; DB 2; Length 1983;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 400 CAAGCTGTCCAAAGTTTAACCTCTGGGAAACCTGTGCTCAAGGAACGACGATGCAAT 459
 Qy 361 GGAATGTCACTGTTCACTTAACTTAAATGTTAAAGATGCACTGTTACATT 420
 Db 460 GGAATGTCACTGTTCACTTAACTTAAATGTTAAAGATGCACTGTTACATT 519
 Qy 421 AAAGAAAGAACCAAAAGAGGTGTAGTCTGCTAGGAATATGTTGGTGGCTTCCAGT 480
 Db 520 AAAGAAAGAACCAAAAGAGGTGTAGTCTGCTAGGAATATGTTGGTGGCTTCCAGT 579
 Qy 481 TACGAATGATCAAGCAACAGATGTTCTTAATAATGGAACAGAAATTAAGCGTT 540
 Db 580 TACGAATGATCAAGCAACAGATGTTCTTAATAATGGAACAGAAATTAAGCGTT 639
 Qy 541 GTTCATATTTATCTTAAATATGTGTAGCAATGATGTTTCAATGTGAAAAAGTA 600
 Db 640 GTTCATATTTATCTTAAATATGTGTAGCAATGATGTTTCAATGTGAAAAAGTA 699
 Qy 601 GGAATGTGAAAAATGAAAGATTAATGAGCGAGAAATTTGTTATCTTAAAGCGAGC 660
 Db 700 GGAATGTGAAAAATGAAAGATTAATGAGCGAGAAATTTGTTATCTTAAAGCGAGC 759
 Qy 661 TCACGAGCAGCAGTAAATATATCCAGAGTCAAGATGATTAATATATGATGACAAAG 720
 Db 760 TCACGAGCAGCAGTAAATATATCCAGAGTCAAGATGATTAATATATGATGACAAAG 819
 Qy 721 GATTAAGAACAGCAAAACGCTTTATTAATGAGAAAGTTATGAATTTGGCAAAATGAT 780
 Db 820 GATTAAGAACAGCAAAACGCTTTATTAATGAGAAAGTTATGAATTTGGCAAAATGAT 879
 Qy 761 TTCAGAGAGCAGAAATGGAACCGGAAATTAACATTAATTAATTAATGAGTTGGTTG 840
 Db 880 TTCAGAGAGCAGAAATGGAACCGGAAATTAACATTAATTAATTAATGAGTTGGTTG 939
 Qy 841 TATATTTTGAAGAGTAAAGTCTCAATATATGAGATTAATTAATTAATTAATTAAT 900
 Db 940 TATATTTTGAAGAGTAAAGTCTCAATATATGAGATTAATTAATTAATTAATTAAT 999
 Qy 901 ACACATTTTACATTTGAGCAACATCAATCAACCTGTTGAAAAAAGTCAAAATGAT 960
 Db 1000 ACACATTTTACATTTGAGCAACATCAATCAACCTGTTGAAAAAAGTCAAAATGAT 1059
 Qy 961 ACCCTTAAGTTGTAATAAACAACACAGCTTAGATGTAATGATGGAATTTGGCGAA 1020
 Db 1060 ACCCTTAAGTTGTAATAAACAACACAGCTTAGATGTAATGATGGAATTTGGCGAA 1119
 Qy 1021 AAATTAATTAATCAAAATTTCTGTAATATTCATTTGGGATTCAGACAAAGAGCGAC 1080
 Db 1120 AAATTAATTAATCAAAATTTCTGTAATATTCATTTGGGATTCAGACAAAGAGCGAC 1179
 Qy 1081 GCTAATTAATTAATCAAAATTTCTGTAATATTCATTTGGGATTCAGACAAAGAGCGAC 1140
 Db 1180 GCTAATTAATTAATCAAAATTTCTGTAATATTCATTTGGGATTCAGACAAAGAGCGAC 1239
 Qy 1141 AACGTGATTTCTGAGAGATATGCTTATGATTAATGATGAGATGATGATGATGAT 1200
 Db 1240 AACGTGATTTCTGAGAGATATGCTTATGATTAATGATGAGATGATGATGATGAT 1299
 Qy 1201 GAAATTAATCAAGTATGTAACAAGCAATGAGCTTCACTGTGCGCTTAATTCAGCGTAT 1260
 Db 1300 GAAATTAATCAAGTATGTAACAAGCAATGAGCTTCACTGTGCGCTTAATTCAGCGTAT 1359
 Qy 1261 ATTCCTAGCTTACACAGCGGCGACACTAAATTCGTTTACTTTATGATTAATGAA 1320
 Db 1360 ATTCCTAGCTTACACAGCGGCGACACTAAATTCGTTTACTTTATGATTAATGAA 1419
 Qy 1321 AAAGAGATTCCTAGAAAGGCTTTAAATGAGGGAATGTTGATATGATGATGATGAT 1380
 Db 1420 AAAGAGATTCCTAGAAAGGCTTTAAATGAGGGAATGTTGATATGATGATGATGAT 1479
 Qy 1381 GACCAAAACCAACCAACCTGTTGAAGTTGTGAACAGTGGGAAACGTTTATTAATGAT 1440
 Db 1480 GACCAAAACCAACCAACCTGTTGAAGTTGTGAACAGTGGGAAACGTTTATTAATGAT 1539

QY 1441 GGGCATGTGACAGACGACAAAGCTTGGCGGAGCTTCCTTGTGTGTCGATGCAAAAC 1500
DB 1540 GGGCATGTGACAGACGACAAAGCTTGGCGGAGCTTCCTTGTGTGTCGATGCAAAAC 1599
QY 1501 AGCGACAGCAAAATTTATTTGAAATCGATGAAACAGAAAGCAGCAACTTGGGTGAAA 1560
DB 1600 AGCGACAGCAAAATTTATTTGAAATCGATGAAACAGAAAGCAGCAACTTGGGTGAAA 1659
QY 1561 ACAAAGCTGAAAGCACTACTTTTACAAACAGCGCTGATGATTAGTTGATATCAAGGG 1620
DB 1660 ACAAAGCTGAAAGCACTACTTTTACAAACAGCGCTGATGATTAGTTGATATCAAGGG 1719
QY 1621 CTAAATACGGTACTATTTATTTAGAAAGAACTGATGCTCCGATGATTATGCTTGTGA 1680
DB 1720 CTAAATACGGTACTATTTATTTAGAAAGAACTGATGCTCCGATGATTATGCTTGTGA 1779
QY 1681 ACAAATCGGATTTGATTTGTGTGTCATGACAAATCATATGCGACAAAGAAACCTAGTT 1740
DB 1780 ACAAATCGGATTTGATTTGTGTGTCATGACAAATCATATGCGACAAAGAAACCTAGTT 1839
QY 1741 TCACCAAGAAAAGTACCAACAAACAAAGTACCTTACCTTCAACAGGTGGCAAGGA 1800
DB 1840 TCACCAAGAAAAGTACCAACAAACAAAGTACCTTACCTTCAACAGGTGGCAAGGA 1899
QY 1801 ATCTACGTTTACTTGAAGTGGCGAGCTTGTCTACTTATTTGACAGAGTCTACTTTGCT 1860
DB 1900 ATCTACGTTTACTTGAAGTGGCGAGCTTGTCTACTTATTTGACAGAGTCTACTTTGCT 1959
QY 1861 AGACGTAGAAAAAAGAAATGCT 1881
DB 1960 AGACGTAGAAAAAAGAAATGCT 1980

RESULT 3
ABN98093
ID ABN98093 standard; DNA; 1983 BP.

XX AC ABN98093;
XX DT 05-AUG-2002 (first entry)
XX DE E faecalis EF058 gene.
XX KW Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic;
XX KM gene; ds.
XX OS Enterococcus faecalis.
XX PN US2002045737-A1.
XX PD 18-APR-2002.
XX PF 04-MAY-1998; 98US-00071035.
XX PR 04-MAY-1998; 98US-00071035.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Choi GH, Bailey C, Hromocky J, Kunsch CA;
XX DR WPI: 2002-425450/45.
XX DR P-P8DB; ABP43337.

PT New genes and polypeptides from Enterococcus faecalis, useful as vaccines
PT for preventing, treating or attenuating an infection caused by a member
PT of the Enterococcus genus in an animal, particularly E. faecalis.
XX
XX PS Claim 1; Page 92; 255pp; English.

CC The present invention provides the protein and coding sequences of a
CC number of polypeptides from Enterococcus faecalis. The proteins can be
CC used as vaccines for preventing or attenuating an infection caused by a

CC member of the Enterococcus genus in an animal, particularly E. faecalis.
CC The polynucleotide is also useful for preventing or treating E. faecalis
CC infection. The present sequence is a coding sequence of the invention
XX
XX SQ Sequence 1983 BP; 729 A; 330 C; 426 G; 498 T; 0 U; 0 Other;

Query Match 100.0%; Score 1881; DB 6; Length 1983;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCAATTAATAAAAGTTGGTACACCGTACTTCTGTATCTAATTTTGGCACTT 60
DB 100 ATGAAGCAATTAATAAAAGTTGGTACACCGTACTTCTGTATCTAATTTTGGCACTT 159
QY 61 TTCACAGTATTTAGGACAAACATGCTGATTTGCAAGAAATGGGAGAGCGCACAG 120
DB 160 TTCACAGTATTTAGGACAAACATGCTGATTTGCAAGAAATGGGAGAGCGCACAG 219
QY 121 CTCGTGATTCACAAAAGAAATGACGAGATTTCACAGATCCGCTTATTCAAAATAGCGG 180
DB 220 CTCGTGATTCACAAAAGAAATGACGAGATTTCACAGATCCGCTTATTCAAAATAGCGG 279
QY 181 AAAGAAATGACGAGATTGATTAATATCAAGGATGCGCATGATGACGTTTATGTTAT 240
DB 280 AAAGAAATGACGAGATTGATTAATATCAAGGATGCGCATGATGACGTTTATGTTAT 339
QY 241 AACGTGAGAAAGCAATTTTACAGACAGAGCGGCGCAAGCGTTGATGACGCTAAA 300
DB 340 AACGTGAGAAAGCAATTTTACAGACAGAGCGGCGCAAGCGTTGATGACGCTAAA 399
QY 301 CAAGCTGTCCAAAGTTTAACTCTCGGAAACCTGTGTCTCAAGAAACACGATGCAAT 360
DB 400 CAAGCTGTCCAAAGTTTAACTCTCGGAAACCTGTGTCTCAAGAAACACGATGCAAT 459
QY 361 GGGAAATGCACTGTTCACTTACCTTAAACCAAAATGTTAAAGATGCTGATATCCATT 420
DB 460 GGGAAATGCACTGTTCACTTACCTTAAACCAAAATGTTAAAGATGCTGATATCCATT 519
QY 421 AAAGAAAGCAAAAGAGGTGTGTGCTGCTGCAATATGATGATGATGATGATGATGAT 480
DB 520 AAAGAAAGCAAAAGAGGTGTGTGCTGCTGCAATATGATGATGATGATGATGATGAT 579
QY 541 GTTCATATTTATCTTAAATAATGTCAGCAATGATGATGATGATGATGATGATGAT 600
DB 580 GTTCATATTTATCTTAAATAATGTCAGCAATGATGATGATGATGATGATGATGAT 699
QY 601 GGAATGCTGAAATGAAAGATTAATGCGCAGAAATTTGTTATTTCTAAAGCGAAGGC 660
DB 700 GGAATGCTGAAATGAAAGATTAATGCGCAGAAATTTGTTATTTCTAAAGCGAAGGC 759
QY 661 TCACCAAGCAGAGTAAATATATCCAGAGTCAAGATGATTAATATATATGACGCAACG 720
DB 760 TCACCAAGCAGAGTAAATATATCCAGAGTCAAGATGATTAATATATATGACGCAACG 819
QY 721 GATTAAGAACAGCAAAAGCTTTATTAATGCGGAAAGATTTATGAATTTGGCAAAATGAT 780
DB 820 GATTAAGAACAGCAAAAGCTTTATTAATGCGGAAAGATTTATGAATTTGGCAAAATGAT 879
QY 781 TTCACAGAGCAGAGTAAATGAAACGGGAAATTAACAGTTAAATCTTGAGGTGCTCG 840
DB 880 TTCACAGAGCAGAGTAAATGAAACGGGAAATTAACAGTTAAATCTTGAGGTGCTCG 939
QY 841 TATATTTAGAGAGTAAAGCTCCAAATATGACGATTAATGAATTCACAAACAAA 900
DB 940 TATATTTAGAGAGTAAAGCTCCAAATATGACGATTAATGAATTCACAAACAAA 999
QY 901 ACACCATTTACATTTGAGCAACCAATCAACACTGTTGAAAAACGTCACAAATATGAT 960
DB 1000 ACACCATTTACATTTGAGCAACCAATCAACACTGTTGAAAAACGTCACAAATATGAT 1059

```

QY 961 ACCTCTAAAGTTGATAAACCAACCAAGCTTAGATGTAAGATGTGGCAATTGGCGAA 1020
DB 1060 ACCTCTAAAGTTGATAAACCAACCAAGCTTAGATGTAAGATGTGGCAATTGGCGAA 1119
QY 1021 AAAATTAATCAAAATTTCTGTAAATATTCATTTGGGGATTTGCAGACAAAGAGCGAC 1080
DB 1120 AAAATTAATCAAAATTTCTGTAAATATTCATTTGGGGATTTGCAGACAAAGAGCGAC 1179
QY 1081 GCTAATTAATAGTCAAAATTTAGTTGATTAACATGATGCAAGCTTAACTTTTGAAT 1140
DB 1180 GCTAATTAATAGTCAAAATTTAGTTGATTAACATGATGCAAGCTTAACTTTTGAAT 1239
QY 1141 AACGTACTCTTGAGAGATGCTTATGCGTTATATGATGGGATACAGTGAATGCTCCT 1200
DB 1240 AACGTACTCTTGAGAGATGCTTATGCGTTATATGATGGGATACAGTGAATGCTCCT 1299
QY 1201 GAAATTTATCAAGTGACTGAACAAAGAAATGCTTCACTGTCGCGCTTAATCCAGCGAT 1260
DB 1300 GAAATTTATCAAGTGACTGAACAAAGAAATGCTTCACTGTCGCGCTTAATCCAGCGAT 1359
QY 1261 ATTCTAGCGCTTAACACCAAGCGGACACCTAAATTTGTTACTTTATGATTAATGA 1320
DB 1360 ATTCTAGCGCTTAACACCAAGCGGACACCTAAATTTGTTACTTTATGATTAATGA 1419
QY 1321 AAAGCAGATCTTACGAAAGGCTTTTAAATTAAGGCGAATGTTGATTAACGCTATCCGAC 1380
DB 1420 AAAGCAGATCTTACGAAAGGCTTTTAAATTAAGGCGAATGTTGATTAACGCTATCCGAC 1479
QY 1381 GACCAAAACCAACCACTGTGTAAGTGTGACAGGTGGGAAAGCTTCAATTAAGTCGAT 1440
DB 1480 GACCAAAACCAACCACTGTGTAAGTGTGACAGGTGGGAAAGCTTCAATTAAGTCGAT 1539
QY 1441 GGGCATGTGACAGCGACACCAACCTTTGGGGAGCTTCTTTGTGTCCGTGATCAAAAC 1500
DB 1540 GGGCATGTGACAGCGACACCAACCTTTGGGGAGCTTCTTTGTGTCCGTGATCAAAAC 1599
QY 1501 AGCGACACGCAAAATTTATTTGAAATTCGATGAAACAAAGAAAGCACTTGGGTGAA 1560
DB 1600 AGCGACACGCAAAATTTATTTGAAATTCGATGAAACAAAGAAAGCACTTGGGTGAA 1659
QY 1561 ACAAAAGCTGAAGCACTACTTTTCAACCAACGGCTGATGATTAAGTATATCAACAGG 1620
DB 1660 ACAAAAGCTGAAGCACTACTTTTCAACCAACGGCTGATGATTAAGTATATCAACAGG 1719
QY 1621 CTTAATTAACGCTTACTTATTTAAGAAACTGTAGCTCCGTGATGATTAAGTATATCA 1680
DB 1720 CTTAATTAACGCTTACTTATTTAAGAAACTGTAGCTCCGTGATGATTAAGTATATCA 1779
QY 1681 ACAATTCGATGAAATTTGTGTCATATGAAACATATGCAACAAAGAAACCTAGTT 1740
DB 1780 ACAATTCGATGAAATTTGTGTCATATGAAACATATGCAACAAAGAAACCTAGTT 1839
QY 1741 TCACCAAGAAAGTACCAAAACCAACCAAGTACTTCACTTCAACAGTGGCAAGGA 1800
DB 1840 TCACCAAGAAAGTACCAAAACCAACCAAGTACTTCACTTCAACAGTGGCAAGGA 1899
QY 1801 ATCTACGTTTACTTGAAGTGGCGCAGCTTGTCTACTTATTTGAGAGTCTACTTTGCT 1860
DB 1900 ATCTACGTTTACTTGAAGTGGCGCAGCTTGTCTACTTATTTGAGAGTCTACTTTGCT 1959
QY 1861 AGACGTAGAAAGAAATGCT 1881
DB 1960 AGACGTAGAAAGAAATGCT 1980

```

```

RESULT 4
ID ACA88057
XX ACA88057 standard; DNA; 1983 BP.
AC ACA88057;
XX 07-JUL-2003 (first entry)

```

```

XX DE E. faecalis novel gene #109.
XX KM Gene; da; endocarditis; bacteremia; urinary tract infection; URT;
XX KM intraabdominal infection; soft tissue infection; neonatal sepsis;
XX KM vaccine.
XX OS Enterococcus faecalis.
XX PN US2003017495-A1.
XX PD 23-JAN-2003.
XX PF 29-JUL-2002; 2002US-00206576.
XX PR 06-MAY-1997; 97US-0044031P.
XX PR 16-MAY-1997; 97US-0046655P.
XX PR 14-NOV-1997; 97US-0066009P.
XX PR 04-MAY-1998; 98US-00071035.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX PT MPI; 2003-416890/39.
XX DR P-PSDB; ABU88365.
XX PT New nucleic acid molecules and polypeptides from Enterococcus faecalis,
XX PT useful as vaccines for preventing or attenuating an enterococcal
XX PT infection in an animal, or for identifying Enterococcus faecalis in
XX PT biological samples.
XX PS Claim 1, Page; 40pp; English.
XX CC The invention relates to a new isolated nucleic acid molecule comprising
XX CC a polynucleotide isolated from Enterococcus faecalis appearing as
XX CC ACA87949-ACA88196 (or sequences complementary to them or 95% identical to
XX CC them). Also included are the proteins encoded by the above nucleic acids,
XX CC making a recombinant vector (comprising inserting the isolated nucleic
XX CC acid molecule cited above into a vector), a host cell comprising the
XX CC vector (used to produce the protein), an isolated antibody specific for
XX CC the polypeptides, a hybridoma that produces the antibody, an isolated
XX CC polypeptide antigen comprising an amino acid sequence of an Enterococcus
XX CC faecalis epitope listed in the specification, a vaccine comprising one or
XX CC more E. faecalis polypeptides (and a pharmaceutical diluent carrier or
XX CC excipient) where the polypeptide elicits protective antibodies in an
XX CC animal to a member of the genus Enterococcus; preventing or attenuating
XX CC an infection caused by a member of the genus Enterococcus in an animal
XX CC comprising administering to the animal the polypeptide and detecting
XX CC Enterococcus nucleic acids in a biological sample. The E. faecalis
XX CC nucleic acid molecules and polypeptides are useful as vaccines for
XX CC preventing or attenuating an enterococcal infection in an animal (e.g.
XX CC endocarditis, bacteraemia, urinary tract infection (URT), intraabdominal
XX CC infection, soft tissue infection and neonatal sepsis). The polypeptides
XX CC are also useful for detecting Enterococcus aureus in immunoassays, as
XX CC epitope tags, as molecular weight markers, or for generating antibodies
XX CC that specifically bind E. faecalis polypeptides. The nucleic acid
XX CC molecules are also useful as probes for gene mapping, or for identifying
XX CC E. faecalis in biological samples. The kit and methods are useful for
XX CC detecting Enterococcus antibodies or nucleic acid molecules in a
XX CC biological sample. The present sequence is a novel E. faecalis nucleic
XX CC acid of the invention. Note: The sequence data for this patent did not
XX CC form part of the printed specification, but was obtained in electronic
XX CC format directly from the USPTO at
XX CC seqdata.uspto.gov/sequence.html?docID=20030017495
XX SQ Sequence 1983 BP; 729 A; 330 C; 426 G; 498 T; 0 U; 0 Other;

```

```

Query Match 100.0%; Score 1881; DB 8; Length 1983;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGCAATTAAAGTTGGTACACCGTTAGTACTTGTACTAATTTTGCACCTT 60

```

Db 100 ATGAGCAATTTAAAAAGTTTGGTACACCGTTAGTACTCTTTACTAATTTTGGCACTT 159
Qy 61 TTCAAGTGTATTAGGGAACAACCTGCATTGGAGAAGAAATGGGAGAGCGCACAG 120
Db 160 TTCAAGTGTATTAGGGAACAACCTGCATTGGAGAAGAAATGGGAGAGCGCACAG 219
Qy 121 CTCGGATTTCACAAAAAGAAATGACGGATTTCACAGATCCGGCTTATTCAAATAGCGGG 180
Db 220 CTCGGATTTCACAAAAAGAAATGACGGATTTCACAGATCCGGCTTATTCAAATAGCGGG 279
Qy 181 AAAAGAAATGAGCGAGTTGTATTAATATCAAGGACTGGCAGATGTGACGTTTATGATTTAT 240
Db 280 AAAAGAAATGAGCGAGTTGTATTAATATCAAGGACTGGCAGATGTGACGTTTATGATTTAT 339
Qy 241 AACGTGACGAACGAATTTTACGAGCAACGAGCGGAGCGGCAACGCGTTGATGACGCTAAA 300
Db 340 AACGTGACGAACGAATTTTACGAGCAACGAGCGGAGCGGCAACGCGTTGATGACGCTAAA 399
Qy 301 CAAGCTGTCCAAAGTTTAACTCTGGGAAAACCTGTTCCTCAAGGAACCAACGATGCAAT 360
Db 400 CAAGCTGTCCAAAGTTTAACTCTGGGAAAACCTGTTCCTCAAGGAACCAACGATGCAAT 459
Qy 361 GGGAAATGTCAGTGTTCAGTTACCTTAAAAAACAATGTAAGATGCAAGTATACAT 420
Db 460 GGGAAATGTCAGTGTTCAGTTACCTTAAAAAACAATGTAAGATGCAAGTATACAT 519
Qy 421 AAAAGAAACCAAAAGAGGGGTAGTGTCTGCTACGATATAGTGTCGCTCCCAT 480
Db 520 AAAAGAAACCAAAAGAGGGGTAGTGTCTGCTACGATATAGTGTCGCTCCCAT 579
Qy 481 TACGAATGATCAAGCAACAGATGGTCTCTTAAATATGGAACGAAGATTAGCGGTT 540
Db 580 TACGAATGATCAAGCAACAGATGGTCTCTTAAATATGGAACGAAGATTAGCGGTT 639
Qy 541 GTTCAATATTTATCTTAAAAATGTGTAGCAATGATGTATGTTTACATGTGAAAAAGTA 600
Db 640 GTTCAATATTTATCTTAAAAATGTGTAGCAATGATGTATGTTTACATGTGAAAAAGTA 699
Qy 601 GGAATGCTGAAAAATGAGAGATTAAATGGCGCAGAAATTTGTTATTTCTTAAACCGAAGGC 660
Db 700 GGAATGCTGAAAAATGAGAGATTAAATGGCGCAGAAATTTGTTATTTCTTAAACCGAAGGC 759
Qy 661 TCACCAAGGACAGTAAATATATCCAGGAGTCAAGATGATGATTAATATCATGAGCAAG 720
Db 760 TCACCAAGGACAGTAAATATATCCAGGAGTCAAGATGATGATTAATATCATGAGCAAG 819
Qy 721 GATTAAGAACCAAGCAAAACGCTTTATTAAGTGGAAAAAGTTATGAAATTTGGCGAAAATGAT 780
Db 820 GATTAAGAACCAAGCAAAACGCTTTATTAAGTGGAAAAAGTTATGAAATTTGGCGAAAATGAT 879
Qy 781 TTCACGAAGACAGAGATGAGAACGGGAGAAATTAACGTTTAAAAATCTTGAAGTGTG 840
Db 880 TTCACGAAGACAGAGATGAGAACGGGAGAAATTAACGTTTAAAAATCTTGAAGTGTG 939
Qy 841 TATATTTTGAAGAGTAAAGGCTCCAATTAATGCAAGATTAATGAATCAACAAAA 900
Db 940 TATATTTTGAAGAGTAAAGGCTCCAATTAATGCAAGATTAATGAATCAACAAAA 999
Qy 901 ACACCAATTTACAATTGAGCAAAACATCAAAACGCTGTGAAAAAAACAGTCAAAAATGAT 960
Db 1000 ACACCAATTTACAATTGAGCAAAACATCAAAACGCTGTGAAAAAAACAGTCAAAAATGAT 1059
Qy 961 ACCTCTAAAGTTGATTAACACACCAAGCTTGAATGATGAAGATGTCGCAATTTGGCGAA 1020
Db 1060 ACCTCTAAAGTTGATTAACACACCAAGCTTGAATGATGAAGATGTCGCAATTTGGCGAA 1119
Qy 1021 AAAATTAATATCAAAATTTCTGTAATATTTCAATGGGGATTCAGACAAAGAGCGCAC 1080
Db 1120 AAAATTAATATCAAAATTTCTGTAATATTTCAATGGGGATTCAGACAAAGAGCGCAC 1179
Qy 1081 GCTAATTAATACGTCAATTTCAATTTAGTTATTAACATGATGACGCTTAACTTTTGAAT 1140

Db 1180 GCTAATTAATACGTCAATTTCAATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 1239
Qy 1141 AACGTGACTTCTGAGAGATATGCTTAATGCGTTTATATGATGGGAGATAGTATGCTCT 1200
Db 1240 AACGTGACTTCTGAGAGATATGCTTAATGCGTTTATATGATGGGAGATAGTATGCTCT 1299
Qy 1201 GAAATTAATCAAGTACTGAAACAAGCAATGAGCTTCACTGTGCGCGTTAATTCAGCGTAT 1260
Db 1300 GAAATTAATCAAGTACTGAAACAAGCAATGAGCTTCACTGTGCGCGTTAATTCAGCGTAT 1359
Qy 1261 ATTCCTAGCTTAACACAGCGGCGACACTAAATTCGTTTACTTTATGATTAATGA 1320
Db 1360 ATTCCTAGCTTAACACAGCGGCGACACTAAATTCGTTTACTTTATGATTAATGA 1419
Qy 1321 AAAGCAATTCCTACGAAGGCTTTAAATAGAGGGAATGTTGATTAAGGTATACCGAC 1380
Db 1420 AAAGCAATTCCTACGAAGGCTTTAAATAGAGGGAATGTTGATTAAGGTATACCGAC 1479
Qy 1381 GACCAAAACACCAACCACTGTTGAAGTTGTGACAGGTGGGAAAACGTTTCAATTAAGTCAT 1440
Db 1480 GACCAAAACACCAACCACTGTTGAAGTTGTGACAGGTGGGAAAACGTTTCAATTAAGTCAT 1539
Qy 1441 GCGGATGTGACAGCAACAAGCTTGGCGGAGCTTCTTTGTGCTCGGTATCAAAAC 1500
Db 1540 GCGGATGTGACAGCAACAAGCTTGGCGGAGCTTCTTTGTGCTCGGTATCAAAAC 1599
Qy 1501 AGCGACAGCAAAATTTATTTGAAATTCATGATGAACAAGAAAGACCAACTGGTGAAA 1560
Db 1600 AGCGACAGCAAAATTTATTTGAAATTCATGATGAACAAGAAAGACCAACTGGTGAAA 1659
Qy 1561 ACAAAAGCTGAAGCAACTATTTTACAACAACGCGCTGATGATTAATGTTATATCAAGG 1620
Db 1660 ACAAAAGCTGAAGCAACTATTTTACAACAACGCGCTGATGATTAATGTTATATCAAGG 1719
Qy 1621 CTTAAATACGTTACTTATTTATTAAGAAACTGATGCTCCGATGATTAATGCTTTGTTA 1680
Db 1720 CTTAAATACGTTACTTATTTATTAAGAAACTGATGCTCCGATGATTAATGCTTTGTTA 1779
Qy 1681 ACAATTCGATTTGAATTTGTGTCATATGAACAATCATATGCGACAACGAAAACTAGTT 1740
Db 1780 ACAATTCGATTTGAATTTGTGTCATATGAACAATCATATGCGACAACGAAAACTAGTT 1839
Qy 1741 TCACCAAGAAAAAGTACCAACCAACAAAGGTAACCTTACCTTCAACAGTGGCAAGGA 1800
Db 1840 TCACCAAGAAAAAGTACCAACCAACAAAGGTAACCTTACCTTCAACAGTGGCAAGGA 1899
Qy 1801 ATCTACGTTTACTTGAAGATGGCGCACTTGTCTTACTTATTTGAGAGTCTTACTTGTCT 1860
Db 1900 ATCTACGTTTACTTGAAGATGGCGCACTTGTCTTACTTATTTGAGAGTCTTACTTGTCT 1959
Qy 1861 AGACGTAGAAAAAGAAATGCT 1881
Db 1960 AGACGTAGAAAAAGAAATGCT 1980

RESULT 5
ABX61663
ID ABX61663 strand: DNA; 1983 BP.
XX
AC ABX61663;
XX
DT 26-FEB-2003 (first entry)
XX
DE Enterococcus faecalis EF040 polymnucleotide #109.
XX
KW EF040; gene; ds; immunostimulant; antibacterial; gene mapping.
XX
OS Enterococcus faecalis.
XX
PN US6448043-B1.
XX
PD 10-SBP-2002.
XX

PF 04-MAY-1998; 98US-00071035.
XX
PR 06-MAY-1997; 97US-0044031P.
PR 16-MAY-1997; 97US-004655P.
PR 14-NOV-1997; 97US-0066009P.
PR 14-NOV-1997; 97US-0066099P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Choi GH, Bailey C, Hromockyj A, Kunsch CA,
XX MPI; 2003-089120/08.
DR P-PSDB; ABUJ3616.
XX
PT New EF040 polypeptides and polynucleotides from *Enterococcus faecalis*,
PT useful for generating an immune response against *E. faecalis* and other
PT *Enterococcus* species, and as vaccines against other bacterial genera.
XX
PS Example 1; Col 113-116; 146p; English.
XX
CC The invention relates to polynucleotide fragments of a gene from
CC *Enterococcus faecalis*, EF040, and the polypeptides encoded by them. The
CC polypeptides are useful in detecting *E. faecalis*, as epitope tags, as
CC molecular weight markers on SDS-PAGE gels or for molecular sieve gel
CC filtration columns, in generating antibodies that specifically bind to
CC the *E. faecalis* polypeptides, in generating an immune response against *E.*
CC *faecalis* and other *Enterococcus* species and as vaccines against other
CC bacterial genera. The polynucleotides are useful as probes for gene
CC mapping and for identifying *E. faecalis* in biological samples. Sequences
CC ABX6155-ABX61802 represent EF040 polynucleotides of the invention. Note:
CC The sequence data for this patent can also be obtained from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 1983 BP; 729 A; 330 C; 426 G; 498 T; 0 U; 0 Other;

Query Match 100.0%; Score 1881; DB 8; Length 1983;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCAATTAATAAAGTTGGTAACCGTTAGTACCTTGTACTTAATTTGGCACTT 60
DB 100 ATGAGCAATTAATAAAGTTGGTAACCGTTAGTACCTTGTACTTAATTTGGCACTT 159
QY 61 TTCACAGGTATTTAGGCAACAACATGCAATTTGCAAGAAATGGGAGACGCAAG 120
DB 160 TTCACAGGTATTTAGGCAACAACATGCAATTTGCAAGAAATGGGAGACGCAAG 219
QY 121 CTCGTGATTCACAAAAGAAAATGACGATTTACAGATCCGCTTATTCAAAATAGCGG 180
DB 220 CTCGTGATTCACAAAAGAAAATGACGATTTACAGATCCGCTTATTCAAAATAGCGG 279
QY 181 AAAAGAAATAGCGAGTTGATTAATAATATCAAGACTGCGAGATGACGTTTACTATTAT 240
DB 280 AAAAGAAATAGCGAGTTGATTAATAATATCAAGACTGCGAGATGACGTTTACTATTAT 339
QY 241 AACGTGACGAAGAAATTTTACAGAGCAACAGCGGCGGCGCAAGCCTTGATGAGCTTAA 300
DB 340 AACGTGACGAAGAAATTTTACAGAGCAACAGCGGCGGCGCAAGCCTTGATGAGCTTAA 399
QY 301 CAAGCTGTCCAAAGTTTAACCTCTGGGAAAACCTGTGTCACAGAGCAACCGATGCAAT 360
DB 400 CAAGCTGTCCAAAGTTTAACCTCTGGGAAAACCTGTGTCACAGAGCAACCGATGCAAT 459
QY 361 GGGAAATGTCAGTGTTCAGTTTAACTTAATAAACAATATGTAAGATGATGATATACATT 420
DB 460 GGGAAATGTCAGTGTTCAGTTTAACTTAATAAACAATATGTAAGATGATGATATACATT 519
QY 421 AAAAGAAAGCAAAAGAGGTGATGCTGCTACGAATATGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 520 AAAAGAAAGCAAAAGAGGTGATGCTGCTACGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 579
QY 481 TAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

DB 580 TAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 639
QY 541 GTTCATATTTTATCTTAAATAATGTGTAGCCATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 640 GTTCATATTTTATCTTAAATAATGTGTAGCCATGATGATGATGATGATGATGATGATGATGATGATGAT 699
QY 601 GGAACCTGCTGAAAATGAAAGATTTAAATGGCGCAAGTTTGTATTTCTAAAGCGAAGCG 660
DB 700 GGAACCTGCTGAAAATGAAAGATTTAAATGGCGCAAGTTTGTATTTCTAAAGCGAAGCG 759
QY 661 TCACGAGGACAGTAAATATATTCACAGAGTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 760 TCACGAGGACAGTAAATATATTCACAGAGTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 819
QY 721 GATTAAGAACAGCAAAAGCCTTATTTACTGGGAAAAGTTATGAAATTTGGCGGAAATGAT 780
DB 820 GATTAAGAACAGCAAAAGCCTTATTTACTGGGAAAAGTTATGAAATTTGGCGGAAATGAT 879
QY 781 TTCACAGAGGACAGTAAATGAAACGGGAGATTTAAGATTTAAATCTTGAGGTTGCTG 840
DB 880 TTCACAGAGGACAGTAAATGAAACGGGAGATTTAAGATTTAAATCTTGAGGTTGCTG 939
QY 841 TATATTTTGAAGAGTAAAGCTCCAAATATATGACAAATTAATGAAATCAACAA 900
DB 940 TATATTTTGAAGAGTAAAGCTCCAAATATATGACAAATTAATGAAATCAACAA 999
QY 901 ACACCTTTTACATTTGAGCAAAATCAATCAACCTGTTGAAAAACAGTCAAAATGAT 960
DB 1000 ACACCTTTTACATTTGAGCAAAATCAATCAACCTGTTGAAAAACAGTCAAAATGAT 1059
QY 961 ACCTTAAAGTTGATTAACCAACACAGACCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1060 ACCTTAAAGTTGATTAACCAACACAGACCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 1119
QY 1021 AAAATTAATATCAAAATTTCTGTAATATTCATTTGGGATTCAGACAAAGAGCGAC 1080
DB 1120 AAAATTAATATCAAAATTTCTGTAATATATTCATTTGGGATTCAGACAAAGAGCGAC 1179
QY 1081 GCTAATTAATATGCTCAAAATTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1180 GCTAATTAATATGCTCAAAATTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1239
QY 1141 AACGTGATTCGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1240 AACGTGATTCGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1299
QY 1201 GAAATTAATCAAGTGACTGAACAGCAATGCGTTCACTGTGCGCTTAATCCAGCGTAT 1260
DB 1300 GAAATTAATCAAGTGACTGAACAGCAATGCGTTCACTGTGCGCTTAATCCAGCGTAT 1359
QY 1261 ATTTCTAGCGTTAACACAGCGCGCACACTTAATTTGTTTACTTTATGCAATTTAAATGAA 1320
DB 1360 ATTTCTAGCGTTAACACAGCGCGCACACTTAATTTGTTTACTTTATGCAATTTAAATGAA 1419
QY 1321 AAAAGAGATCTTACGAAAGGCTTTAAATAATGAGCGAATGTTGATTAACGCTCATACGAC 1380
DB 1420 AAAAGAGATCTTACGAAAGGCTTTAAATAATGAGCGAATGTTGATTAACGCTCATACGAC 1479
QY 1381 GACCAAAACCAACCAACTGTTGAAGTTGACAGGTGGGAAAAGTTTCAATTAAGTCGAT 1440
DB 1480 GACCAAAACCAACCAACTGTTGAAGTTGACAGGTGGGAAAAGTTTCAATTAAGTCGAT 1539
QY 1441 GGGGATGTCACAGGACACAGCCTTGGCGGAGCTTCTTGTGTCCTGCTGATCAAAAC 1500
DB 1540 GGGGATGTCACAGGACACAGCCTTGGCGGAGCTTCTTGTGTCCTGCTGATCAAAAC 1599
QY 1501 AGGACACAGCAAAATTAATTTGAAAATTCATGAAAACAGAAAGCAAGCACTTGGGTGAAA 1560
DB 1600 AGGACACAGCAAAATTAATTTGAAAATTCATGAAAACAGAAAGCAAGCACTTGGGTGAAA 1659
QY 1561 ACAAAAGCTGAAGCAACTTATTTTCAACAAACGCGTATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1660 ACAAAAGCTGAAGCAACTTATTTTCAACAAACGCGTATGATGATGATGATGATGATGATGATGATGAT 1719

QY 1621 CTTAAATAGGCTACTTATTTAGAGAAAGTGTAGCTCTGAGATTTATGTTCTGTA 1680
Db 1720 CTTAAATAGGCTACTTATTTAGAGAAAGTGTAGCTCTGAGATTTATGTTCTGTA 1779
QY 1681 ACAATTCGATTGATTTGTGTCATATGCAATCATATATGCAACAAGAACTAGTT 1740
Db 1780 ACAATTCGATTGATTTGTGTCATATGCAATCATATGCAACAAGAACTAGTT 1839
QY 1741 TCACCAAGAAAGTATCCAAACCAACCAAGTACCTTCAACAGGTGGCAAGGA 1800
Db 1840 TCACCAAGAAAGTATCCAAACCAACCAAGTACCTTCAACAGGTGGCAAGGA 1899
QY 1801 ATCTACGTTACTTGTGAGAGTGGCGAGCTGTCTATCTATGTCAGAGATCTTGTGT 1860
Db 1900 ATCTACGTTACTTGTGAGAGTGGCGAGCTGTCTATCTATGTCAGAGATCTTGTGT 1959
QY 1861 AGACGTAGAGAAAGAAATGCT 1881
Db 1960 AGACGTAGAGAAAGAAATGCT 1980
RESULT 6
ID ADY38951
ADY38951 standard; DNA; 1983 BP.
XX
AC ADY38951;
XX
DT 05-MAY-2005 (first entry)
XX
DE Novel Enterococcus faecalis gene sequence SeqID217.
XX
KM protein purification; DNA purification; antibacterial; vaccine;
XX enterococcus faecalis infection; gene; de.
XX
OS Enterococcus faecalis.
XX
PN US2005043528-A1.
PD 24-FEB-2005.
XX
PF 06-AUG-2004; 2004US-00912362.
XX
PR 06-MAY-1997; 97US-0044031P.
PR 16-NOV-1997; 97US-0046655P.
PR 14-NOV-1997; 97US-0066009P.
PR 04-MAY-1998; 98US-00071035.
PR 29-JUL-2002; 2002US-00206576.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX
DR WPI; 2005-180840/19.
DR P-PSDB; ADY38952.
XX
PT New isolated Enterococcus faecalis nucleic acid useful for diagnosing,
PT preventing and/or attenuating infection by Enterococcus faecalis in
PT animals, particularly in humans.
XX
PS Claim 1; SEQ ID NO 217; 28bp; English.
XX
CC This invention relates to novel genes and the proteins they encode
CC isolated from Enterococcus faecalis. The invention may be useful for the
CC development of compounds with an antibacterial activity or a vaccine. The
CC present invention is useful for the diagnosis, prevention and/or
CC attenuation of infection by Enterococcus faecalis in animals,
CC particularly in humans. The present sequence is that of a novel
CC Enterococcus faecalis gene of the invention. Note: The present sequence
CC does not form part of the printed specification but was obtained in
CC electronic form from the USPTO web site.
XX
SQ Sequence 1983 BP; 729 A; 330 C; 426 G; 498 T; 0 U; 0 Other;

Query Match 100.0%; Score 1881; DB 14; Length 1983;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1881; Conservative 0;
QY 1 ATGAAGCAATTTAAAAAGTTTGTACACCGTTAGTACTTGTATCTAAATTTTGCACCT 60
Db 100 ATGAAGCAATTTAAAAAGTTTGTACACCGTTAGTACTTGTATCTAAATTTTGCACCT 159
QY 61 TTCACAGTGTATTTGGGACAAACCTGATTTTGCAGAAAGAAATGGGAGAGCCGACAG 120
Db 160 TTCACAGTGTATTTGGGACAAACCTGATTTTGCAGAAAGAAATGGGAGAGCCGACAG 219
QY 121 CTCGTGATTTCAAAAAGAAATGACCGATTTACAGATCCGCTTATTTCAAAATAGCGGG 180
Db 220 CTCGTGATTTCAAAAAGAAATGACCGATTTTACAGATCCGCTTATTTCAAAATAGCGGG 279
QY 181 AAGGAATGAGGAGTTTGATTAATATCAAGGACTGGCAGATGTGACGTTTAGTATTTAT 240
Db 280 AAGGAATGAGGAGTTTGATTAATATCAAGGACTGGCAGATGTGACGTTTAGTATTTAT 339
QY 241 AACGTGACGAACGAATTTTACGACGACGAGCGGCAAGCGTTGATGACGCTAA 300
Db 340 AACGTGACGAACGAATTTTACGACGACGAGCGGCAAGCGTTGATGACGCTAA 399
QY 301 CAAGCTGTCCAAAGTTTAACTCTGAGGAAACCTGTTGCTCAAGAAACCAACGATGCAAT 360
Db 400 CAAGCTGTCCAAAGTTTAACTCTGAGGAAACCTGTTGCTCAAGAAACCAACGATGCAAT 459
QY 361 GGGAAATGTCAGTTCAGTTACCTTAACCAAAATGTTAAAGTCAAGTGTATACCAT 420
Db 460 GGGAAATGTCAGTTCAGTTACCTTAACCAAAATGTTAAAGTCAAGTGTATACCAT 519
QY 421 AAAAGAAACCAAAAGAGGTGTAGTGTGTAGGAATATGAGTGTGCGTCCAGTT 480
Db 520 AAAAGAAACCAAAAGAGGTGTAGTGTGTAGGAATATGAGTGTGCGTCCAGTT 579
QY 481 TAGGAATGATCAACCAACAGATGTTCTTAATATGAAACAGAAATTAAGCGTT 540
Db 580 TAGGAATGATCAACCAACAGATGTTCTTAATATGAAACAGAAATTAAGCGTT 639
QY 541 GTTCATATTTATCTTAATAATGTGTAGTGTAGTGTATGTTATGTTATGTTATGTTAT 600
Db 640 GTTCATATTTATCTTAATAATGTGTAGTGTAGTGTATGTTATGTTATGTTATGTTAT 699
QY 601 GGAACCTGTGAATAAGAGATTAAATGCGAGAAATTTGTTATTTCTAAAGCGAAGGC 660
Db 700 GGAACCTGTGAATAAGAGATTAAATGCGAGAAATTTGTTATTTCTAAAGCGAAGGC 759
QY 661 TCACCAAGCAAGTAAATATATCCAGAGTCAAGATGATGATTAATATCATGACCAAG 720
Db 760 TCACCAAGCAAGTAAATATATCCAGAGTCAAGATGATGATTAATATCATGACCAAG 819
QY 721 GATTAAGAAACAGCAAAACGCTTTATTTACTGGGAAAAGTTATGAAATTTGGCAAAATGAT 780
Db 820 GATTAAGAAACAGCAAAACGCTTTATTTACTGGGAAAAGTTATGAAATTTGGCAAAATGAT 879
QY 781 TTCACAGAGCAAGAAATGAGACGGAGAAATTAACGTTAAATCTTGAGGTGGTTGG 840
Db 880 TTCACAGAGCAAGAAATGAGACGGAGAAATTAACGTTAAATCTTGAGGTGGTTGG 939
QY 841 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAATTAATGAAATCAAAACAAA 900
Db 940 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAATTAATGAAATCAAAACAAA 999
QY 901 ACACCAATTTTACATTTGAGCAACCAATCAACCTGTTGAAAACAGTCAAAATGAT 960
Db 1000 ACACCAATTTTACATTTGAGCAACCAATCAACCTGTTGAAAACAGTCAAAATGAT 1059
QY 961 ACCTCTAAAGTGTATTAACCAACCAAGCTTGAATGTTAAAGTGTGCAATTTGGCGAA 1020
Db 1060 ACCTCTAAAGTGTATTAACCAACCAAGCTTGAATGTTAAAGTGTGCAATTTGGCGAA 1119


```

QY 1021 AAAATTAATCAATTTCTGTAATATATTCATTTGGGGATGGAGACAAGAGGCGAC 1080
DB 1120 AAAATTAATCAATTTCTGTAATATATTCATTTGGGGATGGAGACAAGAGGCGAC 1179
QY 1081 GCTAATTAATACGTCAATTTCAATTTAGTTGTAATAACATGACAGCTTAACTTTTGAT 1140
DB 1180 GCTAATTAATACGTCAATTTCAATTTAGTTGTAATAACATGACAGCTTAACTTTTGAT 1239
QY 1141 AACGGAATCTTGGAGAGATATGCTTATATGATGGGGATACAGTATGCTCTCT 1200
DB 1240 AACGGAATCTTGGAGAGATATGCTTATATGATGGGGATACAGTATGCTCTCTCT 1299
QY 1201 GAAAAATTAATCAATTTCAATTTAGTTGTAATAACATGACAGCTTAACTTTTGAT 1260
DB 1300 GAAAAATTAATCAATTTCAATTTAGTTGTAATAACATGACAGCTTAACTTTTGAT 1359
QY 1261 ATTCTAGCTTAACACAGGCGGCACTAAATTTGTTACTTTATGCAATTTAAATGAA 1320
DB 1360 ATTCTAGCTTAACACAGGCGGCACTAAATTTGTTACTTTATGCAATTTAAATGAA 1419
QY 1321 AAAGCAATCTTACGAAAGGCTTTAAATAAGGCGAATGTTGATTAACGCTCATACGAC 1380
DB 1420 AAAGCAATCTTACGAAAGGCTTTAAATAAGGCGAATGTTGATTAACGCTCATACGAC 1479
QY 1381 GACCAAAACGCACTGTTGAGTGTGACAGGTTGGGAAAGCTTCAATTAAGTCGAT 1440
DB 1480 GACCAAAACGCACTGTTGAGTGTGACAGGTTGGGAAAGCTTCAATTAAGTCGAT 1539
QY 1441 GGGCGATGTGACAGCAACAGCCTTGGGCGAGCTTCCCTTGGCTCCGATCAAAAC 1500
DB 1540 GGGCGATGTGACAGCAACAGCCTTGGGCGAGCTTCCCTTGGCTCCGATCAAAAC 1599
QY 1501 AGCGACACGCAATTTATTTGAAATTCATGAAACAGCAAGAGCAACTTGGGTGAA 1560
DB 1600 AGCGACACGCAATTTATTTGAAATTCATGAAACAGCAAGAGCAACTTGGGTGAA 1659
QY 1561 ACAAAAGCTGAAGCACTTCTTCAACAGGCGTGAATGATTTGATATCAACAGG 1620
DB 1660 ACAAAAGCTGAAGCACTTCTTCAACAGGCGTGAATGATTTGATATCAACAGG 1719
QY 1621 CTTAAATACGTAATTTATTTAGAAAGAACTGAGCTCTGATGATTTATCTTTGTA 1680
DB 1720 CTTAAATACGTAATTTATTTAGAAAGAACTGAGCTCTGATGATTTATCTTTGTA 1779
QY 1681 ACAATTCGATTTGATTTGTGCTCATATGAAACATATGCGACAAAGAAACCTAGTT 1740
DB 1780 ACAATTCGATTTGATTTGTGCTCATATGAAACATATGCGACAAAGAAACCTAGTT 1839
QY 1741 TCACAGAAAGAAAGTACCAAGAAACAAAGGTAACCTTCAACAGGTTGCAAAAGGA 1800
DB 1840 TCACAGAAAGAAAGTACCAAGAAACAAAGGTAACCTTCAACAGGTTGCAAAAGGA 1899
QY 1801 ATCTAGTTTACTTGAAGAGTGGCAGCTTGTACTTACTTATGACAGAGTCTACTTGT 1860
DB 1900 ATCTAGTTTACTTGAAGAGTGGCAGCTTGTACTTACTTATGACAGAGTCTACTTGT 1959
QY 1861 AGACGTAGAAAGAAATGCT 1881
DB 1960 AGACGTAGAAAGAAATGCT 1980

```

RESULT 7

AAK12982
ID AAK12982 standard; DNA; 15614 BP.

AAK12982;

19-MAR-1999 (first entry)

Enterococcus faecalis genome contig SEQ ID NO:45.

Enterococcus faecalis; contig; detection; Enterococcal infection;

vacine; attenuation; computer readable medium; ds.

```

XX XX Enterococcus faecalis.
OS WO980555-A2.
PN 12-NOV-1998.
PD 04-MAY-1998; 98WO-US008985.
PR 06-MAY-1997; 97US-0044031P.
PR 16-MAY-1997; 97US-0046555P.
PR 14-NOV-1997; 97US-006009P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Kunesh CA, Dillon PV, Barash SC,
PI WPI; 1999-045171/04.
XX
XX New isolated Enterococcus faecalis polynucleotides and polypeptides -
PT used to develop products for the detection of Enterococcus and for use in
XX vaccines for prevention or attenuation of Enterococcus infection.
XX
XX Claim 1; Page 419-427; 2084pp; English.
XX
XX A computer readable medium has been developed which has recorded on it
XX 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
XX AAK12982 to AAK13919 represent these nucleotide sequences which are
XX primary nucleotide sequences, also known as contigs. The computer-based
XX system can identify fragments of the Enterococcus faecalis genome with
XX commercial importance. The products can be used to detect the presence of
XX Enterococcus faecalis in samples. They can also be used for diagnosing
XX Enterococcal infection in an animal and monitoring progression of
XX disease, and for identifying agents which can be used to modulate the
XX growth or pathogenicity of Enterococcus faecalis, or another related
XX organism, in vivo or in vitro. In particular the polypeptides encoded by
XX the Enterococcus faecalis nucleotide sequences can be used in vaccines to
XX prevent or attenuate an Enterococcal infection
XX
XX Sequence 15614 BP; 5535 A; 2567 C; 3169 G; 4326 T; 0 U; 17 Other;
SQ
Query Match 100.0%; Score 1881; DB 2; Length 15614;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACTTGTACTTAATTTGGCACTT 60
DB 12164 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACTTGTACTTAATTTGGCACTT 12223
QY 61 TTCACAACTGTTTGGGACAACTGCTATTTGCAAGAAATGGGGAGCGCACAG 120
DB 12224 TTCACAACTGTTTGGGACAACTGCTATTTGCAAGAAATGGGGAGCGCACAG 12283
QY 121 CTGCGATTTCAAAAAGAAATGACGGATTTACAGATTCGCTTATTTCAAAATAGCGGG 180
DB 12284 CTGCGATTTCAAAAAGAAATGACGGATTTACAGATTCGCTTATTTCAAAATAGCGGG 12343
QY 181 AAAAGAAATGAGGAGTTGATTAATATATCAAGGACTGGAGATGTGACGTTTATTTAT 240
DB 12344 AAAAGAAATGAGGAGTTGATTAATATATCAAGGACTGGAGATGTGACGTTTATTTAT 12403
QY 241 AACGTGACGAAGCAATTTTACAGACAGAGCGGACAGGCGCAAGGCTTATGACACTAA 300
DB 12404 AACGTGACGAAGCAATTTTACAGACAGAGCGGACAGGCGCAAGGCTTATGACACTAA 12463
QY 301 CAAGCTGTCCAAAGTTTAACTCTCTGGGAAACCTGTGTCTCAAGGAACCAAGATCAAA 360
DB 12464 CAAGCTGTCCAAAGTTTAACTCTCTGGGAAACCTGTGTCTCAAGGAACCAAGATCAAA 12523
QY 361 GGGAAATGTCACCTTCACTTCACTTAATAAACAACAAATGTTAAAGATGACAGTATACAT 420
DB 12524 GGGAAATGTCACCTTCACTTCACTTAATAAACAACAAATGTTAAAGATGACAGTATACAT 12583

```

QY	421	AAAAGAAACCAAAAGGGGTGTAAGTTGCTGCTACGAATATGTTGGCTTCCAGTT	480
Db	12584	AAAAGAAACCAAAAGGGGTGTAAGTTGCTGCTACGAATATGTTGGCTTCCAGTT	12643
QY	481	TACGAATAATATCAAGCAACAGATGGTTCCATTAATATGTGAACAGAAATTTGCCGTT	540
Db	12644	TACGAATAATATCAAGCAACAGATGGTTCCATTAATATGTGAACAGAAATTTGCCGTT	12703
QY	541	GTTCAATTTTATCTTAAAAATGTGTGACCAATGATGGTATTACATGTAAAAAAGTA	600
Db	12704	GTTCAATTTTATCTTAAAAATGTGTGACCAATGATGGTATTACATGTAAAAAAGTA	12763
QY	601	GGAACTGCTGAAAATGAGGATTAATATGGCGCAGATTGTTATTTCTAAAAAGCAGGC	660
Db	12764	GGAACTGCTGAAAATGAGGATTAATATGGCGCAGATTGTTATTTCTAAAAAGCAGGC	12823
QY	661	TCACCAGGCAAGATTAATATATCCAGAGGTCAAAGTGGATTTATATCAATGGCAACG	720
Db	12824	TCACCAGGCAAGATTAATATATCCAGAGGTCAAAGTGGATTTATATCAATGGCAACG	12883
QY	721	GATTAAGAACAAACCAAAACGCTTTATCTGAGAAAAGTTATGAAATTTGGCAAAATGAT	780
Db	12884	GATTAAGAACAAACCAAAACGCTTTATCTGAGAAAAGTTATGAAATTTGGCAAAATGAT	12943
QY	781	TTCCAGAACAGAGAAATGGAACGGAGAAATTAACAGTTAAAAATCTTGAGTGTGTCG	840
Db	12944	TTCCAGAACAGAGAAATGGAACGGAGAAATTAACAGTTAAAAATCTTGAGTGTGTCG	13003
QY	841	TATATTTTGAAGAAAGTAAAGCTCCCAATTAATGACAGATTAAATTGAATTCAAACAAA	900
Db	13004	TATATTTTGAAGAAAGTAAAGCTCCCAATTAATGACAGATTAAATTGAATTCAAACAAA	13063
QY	901	ACACCAATTTTCAATTTGAAGCAACAATTCAAACACCTGTGTAAGAAAAACGTCAAAATGAT	960
Db	13064	ACACCAATTTTCAATTTGAAGCAACAATTCAAACACCTGTGTAAGAAAAACGTCAAAATGAT	13123
QY	961	ACCTCTAAAGTGAATTAACAACAACCAAGCTTAAGTGTAAAGATGTGGCAATTTGGCA	1020
Db	13124	ACCTCTAAAGTGAATTAACAACAACCAAGCTTAAGTGTAAAGATGTGGCAATTTGGCA	13183
QY	1021	AAAAATTAATATCAAAATTTCTGTAAATATTCCAATTTGGGATTTGCAACAAAGAGCGAC	1080
Db	13184	AAAAATTAATATCAAAATTTCTGTAAATATTCCAATTTGGGATTTGCAACAAAGAGCGAC	13243
QY	1081	GCTAATTAATACGTCAAATTCATTTATGTTGATTAACAATGACGCTTAACCTTTGAT	1140
Db	13244	GCTAATTAATACGTCAAATTCATTTATGTTGATTAACAATGACGCTTAACCTTTGAT	13303
QY	1141	AACGTGACTTCTGAGAGATATGCTTAATGCTTAATATGATGGGATACAGTATTCCTCT	1200
Db	13304	AACGTGACTTCTGAGAGATATGCTTAATGCTTAATATGATGGGATACAGTATTCCTCT	13363
QY	1201	GAAAAATTAATGAATGATGAACAAGCAATATGGCTTCACTGTGCGCGTTAATCCAGCGAT	1260
Db	13364	GAAAAATTAATGAATGATGAACAAGCAATATGGCTTCACTGTGCGCGTTAATCCAGCGAT	13423
QY	1261	ATTGCTACGCTAACACCAAGGCGGACACTAAATTTGTTTACTTTATGATTTAATGAA	1320
Db	13424	ATTGCTACGCTAACACCAAGGCGGACACTAAATTTGTTTACTTTATGATTTAATGAA	13483
QY	1321	AAAGCAGATCTTACGAAAGGCTTTAAAAATGAGGGAATGTTGATACGGTCATACGAC	1380
Db	13484	AAAGCAGATCTTACGAAAGGCTTTAAAAATGAGGGAATGTTGATACGGTCATACGAC	13543
QY	1381	GACCAAAACCAACCAACTGTTGAAGTTGTGACAGGTGGGAAACGTTTCAATTAAATCGAT	1440
Db	13544	GACCAAAACCAACCAACTGTTGAAGTTGTGACAGGTGGGAAACGTTTCAATTAAATCGAT	13603
QY	1441	GGGCAATGTGACAGGACCAAGCCTTGGCGGAGCTTCTTTGTTGTCGTTATCAAAAC	1500
Db	13604	GGGCAATGTGACAGGACCAAGCCTTGGCGGAGCTTCTTTGTTGTCGTTATCAAAAC	13663
QY	1501	AGCGACACAGCAAAATTAATTTGAAATATGATGAACAAACGAAGCAGCAATTTGGGTGAA	1560

Db	13664	AGCAGCAGCAAAATTATTTGAAATCGATGAAACAAGACGACCTTGGGTGAAA	13722
Oy	1561	ACAAAGCTGAGCACTACTCTTTTACAAACAACGGCTGATGATTTGAATATCACAGGG	1620
Db	13724	ACAAAGCTGAGCACTACTCTTTTACAAACAACGGCTGATGATTTGAATATCACAGGG	13782
Oy	1621	CTTAATACGGTACTTATTTATTTAGAGAAACTGTAGCTCTTGATGATTTATGCTTGTTA	1680
Db	13784	CTTAATACGGTACTTATTTATTTAGAGAAACTGTAGCTCTTGATGATTTATGCTTGTTA	13842
Oy	1681	ACAAATGGATTTGAATTTGTGTCAATGAAACAATGACATATGAGCAACAAGAAACTAGTT	1740
Db	13844	ACAAATGGATTTGAATTTGTGTCAATGAAACAATGACATATGAGCAACAAGAAACTAGTT	13902
Oy	1741	TCACCAGAAAAAGTACCAACAACAACAAGGTACTTACCTTCAACAGGTGGCAAGA	1800
Db	13904	TCACCAGAAAAAGTACCAACAACAACAAGGTACTTACCTTCAACAGGTGGCAAGA	13962
Oy	1801	ATCTACGTTTACTTAGGAAGTGGCGCAGTCTTGCTACTTATTTGACAGAGTCTACTT	1860
Db	13964	ATCTACGTTTACTTAGGAAGTGGCGCAGTCTTGCTACTTATTTGACAGAGTCTACTT	14022
Oy	1861	AGACGTAGAAAAAGAAATGCT 1881	
Db	14024	AGACGTAGAAAAAGAAATGCT 14044	
RESULT 8			
ABS98777			
ID	ABS98777	standard; DNA, 15614 BP.	
AC	ABS98777;		
XX	18-DEC-2002	(first entry)	
DT	Enterococcus faecalis contig sequence #45.		
XX	Computer readable medium; Enterococcus faecalis; microbe; growth;		
KW	pathogenicity; vaccine; resistance; Enterococcal infection; commercial;		
KW	therapeutic; industrial; fermenting; sugar source; metabolite; vaccine;		
KW	biochip technology; antibacterial; modulator of nucleic acid expression;		
KW	contig; ds.		
OS	Enterococcus faecalis.		
XX	US2002120116-A1.		
PN	29-AUG-2002.		
XX	04-MAY-1998;	98US-00070927.	
PF	04-MAY-1998;	98US-00070927.	
XX	04-MAY-1998;	98US-00070927.	
XX	(KUNSCH C A.		
PA	(DILLON P J.		
PA	(BARASH S.		
XX	Kunsch CA, Dillon PJ, Barash S;		
PI	WPI, 2002-750065/81.		
XX	Computer readable medium having recorded on it a Enterococcus faecalis		
PT	nucleotide sequence useful for detecting diseases related to Enterococcus		
XX	infections in animals.		
XX	Claim 1, Page; 119pp; English.		
XX	The present invention relates to a new computer readable medium with an		
CC	Enterococcus faecalis nucleotide sequence. The invention is useful to		
CC	diagnose the presence of E.faecalis in a sample or determining the		
CC	presence of a specific microbe in a sample. The invention is also useful		
CC	for modulating the growth or pathogenicity of E.faecalis, in a vaccine to		

CC confer resistance to Enterococcal infection, for commercial, therapeutic
CC and industrial purposes, and for fermenting a particular sugar source or
CC to produce a particular metabolite. The invention is useful for detecting
CC diseases related to Enterococcus infections in animals, and for detecting
CC E. faecalis using bioclip technology. The present nucleic acid sequence
CC represents an Enterococcus faecalis contig DNA sequence of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at http://seqdata.uspto.gov

XX Sequence 15614 BP, 5535 A, 2567 C, 3169 G, 4326 T, 0 U, 17 Other;

Query Match 100.0%; Score 1881; DB 6; Length 15614;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTGTGTACTTAATTTTGCACCTT 60
DB 12164 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTGTGTACTTAATTTTGCACCTT 12223
QY 61 TTCAACAAGTGTATTAGGGAACAACACTGATTTGCAAGAAATAAGGGAAGCGCAAG 120
DB 12224 TTCAACAAGTGTATTAGGGAACAACACTGATTTGCAAGAAATAAGGGAAGCGCAAG 12283
QY 121 CTGCTGATTTCAAAAAGAAATGAAGGATTTACAGATCCGCTTATTCAAAATAGCGGG 180
DB 12284 CTGCTGATTTCAAAAAGAAATGAAGGATTTACAGATCCGCTTATTCAAAATAGCGGG 12343
QY 181 AAAAAGATGAGCGAGTTGATTAATATCAAGACCTGGCAGATGTGACGTTTAGTATTAT 240
DB 12344 AAAAAGATGAGCGAGTTGATTAATATCAAGACCTGGCAGATGTGACGTTTAGTATTAT 12403
QY 241 AACGTGACGAAGAANTTTTACGAGCAAGAGCGGCGAGCGCAAGCTTGATGACGTAA 300
DB 12404 AACGTGACGAAGAANTTTTACGAGCAAGAGCGGCGAGCGCAAGCTTGATGACGTAA 12463
QY 301 CAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGTGCTCAAGGAACAACGATGCAAT 360
DB 12464 CAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGTGCTCAAGGAACAACGATGCAAT 12523
QY 361 GGGAAATGTCAGTGTTCAGTTACCTTAATAAACAATGATGATGATGATGATGAT 420
DB 12524 GGGAAATGTCAGTGTTCAGTTACCTTAATAAACAATGATGATGATGATGATGAT 12583
QY 421 AAAAGAAACCAAAAAGAGGTGTAGTGTGCTGACGAATATGTGTGTGCTTCCAGTT 480
DB 12584 AAAAGAAACCAAAAAGAGGTGTAGTGTGCTGACGAATATGTGTGTGCTTCCAGTT 12643
QY 481 TACGAATATGATCAAGCAACAGATGCTCTTAAATATGAAACGAAGATTAAGCGTT 540
DB 12644 TACGAATATGATCAAGCAACAGATGCTCTTAAATATGAAACGAAGATTAAGCGTT 12703
QY 541 GTTCATATTTATCTTAAAAATGTGTGACCAATGATGATGATGATGATGATGAT 600
DB 12704 GTTCATATTTATCTTAAAAATGTGTGACCAATGATGATGATGATGATGATGAT 12763
QY 601 GGAATCTGCGAATAAAGAGATTTAAATGGCGCAAAATTTGTTTCTTAAAGCGAAGC 660
DB 12764 GGAATCTGCGAATAAAGAGATTTAAATGGCGCAAAATTTGTTTCTTAAAGCGAAGC 12823
QY 661 TCACCAAGCAAGTAAATATATCCAAAGAGTCAAGATGATTTATATCATGGAACAAG 720
DB 12824 TCACCAAGCAAGTAAATATATCCAAAGAGTCAAGATGATTTATATCATGGAACAAG 12883
QY 721 GATTAAGAACAGCAAAACGCTTTATATCTGAGAAAAGTTATGAATTTGGCAAAATGAT 780
DB 12884 GATTAAGAACAGCAAAACGCTTTATATCTGAGAAAAGTTATGAATTTGGCAAAATGAT 12943
QY 781 TTCAAGAAAGCAAGATGGAACGGGAAATTAACGTTAAATCTTGAGGTTGGTTCG 840
DB 12944 TTCAAGAAAGCAAGATGGAACGGGAAATTAACGTTAAATCTTGAGGTTGGTTCG 13003
QY 841 TATATTTTGAAGAAATGAAGCTCCAAATATGCAATTAATTAATGAATCAAAACAAA 900
```

```
DB 13004 TATATTTTGAAGAAAGTAAAGCTCCAAATATGCAATTAATTAATGAATCAAAACAAA 13063
QY 901 ACACCAATTTACATTTAGAACCAACATCAACACCTGTTGAAAAAACAGTCAAAATGAT 960
DB 13064 ACACCAATTTACATTTAGAACCAACATCAACACCTGTTGAAAAAACAGTCAAAATGAT 13123
QY 961 ACCTCTAAGTTGATTAACCAACACACACTTATGATGATGATGATGATGATGATGAT 1020
DB 13124 ACCTCTAAGTTGATTAACCAACACACACTTATGATGATGATGATGATGATGATGAT 13183
QY 1021 AAAATTAATATCAAAATTTCTGTAATATTTCCATTTGGGATTCAGACAAAGAGCGAC 1080
DB 13184 AAAATTAATATCAAAATTTCTGTAATATTTCCATTTGGGATTCAGACAAAGAGCGAC 13243
QY 1081 GCTAATTAATACGTCAATTAATTAATGATGATGATGATGATGATGATGATGAT 1140
DB 13244 GCTAATTAATACGTCAATTAATTAATGATGATGATGATGATGATGATGATGAT 13303
QY 1141 AACGTGACCTTGGAGAGTATGCTTATGCTTATATGATGATGATGATGATGATGAT 1200
DB 13304 AACGTGACCTTGGAGAGTATGCTTATGCTTATATGATGATGATGATGATGATGAT 13363
QY 1201 GAAATTTATCAAGTGACTGAACAAGCAATGCTGCTGCGCTTATATCAAGCTAT 1260
DB 13364 GAAATTTATCAAGTGACTGAACAAGCAATGCTGCTGCGCTTATATCAAGCTAT 13423
QY 1261 ATTCCTAGCTTAACACAGCGGCGACACTAAATTTGTTTACTTTATGCAATTTAAATGA 1320
DB 13424 ATTCCTAGCTTAACACAGCGGCGACACTAAATTTGTTTACTTTATGCAATTTAAATGA 13483
QY 1321 AAAGCAATCTTACGAAGAGCTTTTAAATATGAGGGAATGTTGATTAACGCTATACCGAC 1380
DB 13484 AAAGCAATCTTACGAAGAGCTTTTAAATATGAGGGAATGTTGATTAACGCTATACCGAC 13543
QY 1381 GACCAAAACACACCAACTGTTGAATGTTGAACAGGTGGGAAACGTTTCAATTAAGTCAT 1440
DB 13544 GACCAAAACACACCAACTGTTGAATGTTGAACAGGTGGGAAACGTTTCAATTAAGTCAT 13603
QY 1441 GCGAGTGTGACAGCAACACAGCCTTGGCGGAGCTTCTTTGTGCTGCGTGAATCAAAAC 1500
DB 13604 GCGAGTGTGACAGCAACACAGCCTTGGCGGAGCTTCTTTGTGCTGCGTGAATCAAAAC 13663
QY 1501 AGCGACACAGCAAAATTTTGAATATCGATGAAACAAAGCAAGCAACTTGGGTGAA 1560
DB 13664 AGCGACACAGCAAAATTTTGAATATCGATGAAACAAAGCAAGCAACTTGGGTGAA 13723
QY 1561 ACAAAAGCTGAAGCAACTTCTTTCAACACAGGCTGATGATGATGATGATGATGAT 1620
DB 13724 ACAAAAGCTGAAGCAACTTCTTTCAACACAGGCTGATGATGATGATGATGATGAT 13783
QY 1621 CTTAAATACGGTACTTATTTTGAAGAAACCTGATGCTCTGATGATTTATGCTTTGTA 1680
DB 13784 CTTAAATACGGTACTTATTTTGAAGAAACCTGATGCTCTGATGATTTATGCTTTGTA 13843
QY 1681 ACAATTCGATTTGATTTTGTGTCAATGAACATATGGAACAAACCAACCAACCTTACT 1740
DB 13844 ACAATTCGATTTGATTTTGTGTCAATGAACATATGGAACAAACCAACCAACCTTACT 13903
QY 1741 TCACCAAGAAAGTACCAACCAACCAAGGATCCTTACCTTCAACAGGTGGCAAGGA 1800
DB 13904 TCACCAAGAAAGTACCAACCAACCAAGGATCCTTACCTTCAACAGGTGGCAAGGA 13963
QY 1801 ATCTAGCTTTACTTGAAGAGTGGCGAGCTTGTCTACTTATTTGACAGAGTCTTACTTGT 1860
DB 13964 ATCTAGCTTTACTTGAAGAGTGGCGAGCTTGTCTACTTATTTGACAGAGTCTTACTTGT 14023
QY 1861 AGAGTGAAGAAAGAAATGCT 1881
DB 14024 AGAGTGAAGAAAGAAATGCT 14044
```

RESULT 9

AD084888
ID AD084888 standard; DNA; 1875 BP.
XX
AC AD084888;
XX
DT 29-JUL-2004 (first entry)
XX
DE E faecalis surface anchored LPXTG protein gene SegID12.
XX
XX LPXTG; cell wall-anchored surface protein; Gram positive bacterium;
XX extracellular matrix molecule; sequence database; C-terminal;
XX immunoglobulin-like fold region; Ig-like fold region; antibacterial;
XX vaccine; gene therapy; infection; medical device; prosthesis;
XX premature newborn; AIDS; debilitated cancer; bone marrow transplantation;
XX gene; ds.
OS Enterococcus faecalis.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..1875
XX /tag= a
XX /product= "E faecalis surface anchored LPXTG protein"
XX /partial
XX /note= "No stop codon"
XX
XX MO2004025416-A2.
XX
XX PD 25-MAR-2004.
XX
XX PF 15-SEP-2003; 2003WO-US028789.
XX
XX PR 13-SEP-2002; 2002US-0410303P.
XX
XX PA (TEXA) UNIV TEXAS A & M SYSTEM.
XX (INHI-) INHIBITEX INC.
XX PA (UABR-) UAB RES FOUND.
XX
XX PI Hook M, Xu Y, Sillanpaa JV, Sthanam N, Pomuraj K, Patti JM;
PI Hutchins JT, Hall A;
XX
XX WPI: 2004-315684/29.
XX DR P-PSDB; AD084889.
XX
XX
XX Identifying LPXTG-containing cell wall-anchored surface proteins from
XX Gram positive bacteria, for treating infection caused by the bacteria,
XX comprises searching sequence information database for the sequence having
XX LPXTG-motif.
XX
XX PS Claim 18; SEQ ID NO 12; 96bp; English.
XX
XX This invention relates to a novel method of identifying LPXTG-containing
XX cell wall-anchored surface proteins from Gram positive bacteria that bind
XX to an extracellular matrix molecule which comprises searching a database
XX of sequence information for a putative protein sequence having the LPXTG-
XX motif in its C-terminal region and analysing the sequence for the
XX presence of one or more immunoglobulin (Ig)-like fold regions. The
XX invention may be useful for the production of compounds with an
XX antibacterial activity or for production of a vaccine. In addition the
XX disclosed sequences may be useful for gene therapy. The antibody is
XX useful for treating or preventing an infection of Gram-positive bacteria
XX in a human or animal patient. The method and the proteins are useful in
XX generating antibodies for treating and preventing the spread of
XX infections of Gram positive bacteria, for interfering with, or inhibiting
XX binding interactions by Gram positive bacteria, for monitoring the level
XX of gram positive bacterial antigens, or antibodies recognizing the
XX antigens in a human or animal patients suspected of containing the
XX antigens or antibodies, in preventing or reducing infection of medical
XX devices and prostheses caused by such organisms, and in treating or
XX preventing infections in highly susceptible groups such as premature
XX newborns, AIDS and debilitated cancer patients, and bone marrow
XX transplantation. The present sequence is that of a gene which encodes a
XX surface anchored LPXTG protein identified using the method of the
XX invention.

XX
SQ Sequence 1875 BP; 691 A; 317 C; 400 G; 467 T; 0 U; 0 Other;
Query Match 98.8%; Score 1859; DB 12; Length 1875;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1875; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
QY 1 ATGAAGCAATTAAAAAGTTTGTGTACACCGGTAGTACCTTTGTATATTTGGCACTT 60
DB 1 ATGAAGCAATTAAAAAGTTTGTGTACACCGGTAGTACCTTTGTATATTTGGCACTT 60
QY 61 TTCAAGATGTATTTAGGACAAACAACCTGATTTTGCAGAAAGAAATGGGAGGCGACAG 120
DB 61 TTCAAGATGTATTTAGGACAAACAACCTGATTTTGCAGAAAGAAATGGGAGGCGACAG 120
QY 121 CTCGTGATTCACAAAAGAAAATGACCGATTTTACAGATCCGCTTATTCAAAATAGCGGG 180
DB 121 CTCGTGATTCACAAAAGAAAATGACCGATTTTACAGATCCGCTTATTCAAAATAGCGGG 180
QY 181 AAAGAAATGAGCGAGTTTGTATTAATATCAAGACTGGCGAGATGTGACGTTTGTATTTAT 240
DB 181 AAAGAAATGAGCGAGTTTGTATTAATATCAAGACTGGCGAGATGTGACGTTTGTATTTAT 240
QY 241 AACGTGACGAACGAATTTTACGAGCAACGACGCGGACGCGAAGCGTTGATGACACTAA 300
DB 241 AACGTGACGAACGAATTTTACGAGCAACGACGCGGACGCGAAGCGTTGATGACACTAA 300
QY 301 CAAGCTGTCCAAAGTTTAACTCTCTGGGAAACCTGTGTCTCAAGGAAACCAACGATGCAAT 360
DB 301 CAAGCTGTCCAAAGTTTAACTCTCTGGGAAACCTGTGTCTCAAGGAAACCAACGATGCAAT 360
QY 361 GGGAAATGTCATCTGTTCAGTTTACCTTAAACCAAAATGGTAAAGATGCAATGATTCATT 420
DB 361 GGGAAATGTCATCTGTTCAGTTTACCTTAAACCAAAATGGTAAAGATGCAATGATTCATT 420
QY 421 AAAAGAAACCAAAAGAGGTGTAGTTCCTGCTACGAATATGTGTGTGCGTTCCAGTT 480
DB 421 AAAAGAAACCAAAAGAGGTGTAGTTCCTGCTACGAATATGTGTGTGCGTTCCAGTT 480
QY 481 TACGAATATGATTAACCAACAGATGTCTTAAATATGAAACGAAGAAATTAACGGTT 540
DB 481 TACGAATATGATTAACCAACAGATGTCTTAAATATGAAACGAAGAAATTAACGGTT 540
QY 541 GTTCATATTTATCTTAAATATGTGTAGCGAATGTATGATGTTTCAATGTGAAAAAGTA 600
DB 541 GTTCATATTTATCTTAAATATGTGTAGCGAATGTATGATGTTTCAATGTGAAAAAGTA 600
QY 594 GGAACCTGCTGAAAATGAAAGATTAATGCGCAGAAATTTGTTATTTCTTAAACCGAAGGC 660
DB 594 GGAACCTGCTGAAAATGAAAGATTAATGCGCAGAAATTTGTTATTTCTTAAACCGAAGGC 660
QY 661 TCACGACGACAGTAAATATATTCAGAGAGTCAGAAAGATTTATATATATGAGCAACG 720
DB 661 TCACGACGACAGTAAATATATTCAGAGAGTCAGAAAGATTTATATATATGAGCAACG 720
QY 714 TCACGACGACAGTAAATATATTCAGAGAGTCAGAAAGATTTATATATATGAGCAACG 780
DB 714 TCACGACGACAGTAAATATATTCAGAGAGTCAGAAAGATTTATATATATGAGCAACG 780
QY 721 GATAAAGAACCAAGCAAAACGCTTTTATCTGCGAAAAGTTTGAATTTGGCGAAATGAT 840
DB 721 GATAAAGAACCAAGCAAAACGCTTTTATCTGCGAAAAGTTTGAATTTGGCGAAATGAT 840
QY 774 GATTAAGAACCAAGCAAAACGCTTTTATCTGCGAAAAGTTTGAATTTGGCGAAATGAT 840
DB 774 GATTAAGAACCAAGCAAAACGCTTTTATCTGCGAAAAGTTTGAATTTGGCGAAATGAT 840
QY 841 TATATTTTGAAGAAAGTAAAGCTTCAATTAATGAGAAATTAATGAAATCAACAAA 900
DB 841 TATATTTTGAAGAAAGTAAAGCTTCAATTAATGAGAAATTAATGAAATCAACAAA 900
QY 894 TATATTTTGAAGAAAGTAAAGCTTCAATTAATGAGAAATTAATGAAATCAACAAA 960
DB 894 TATATTTTGAAGAAAGTAAAGCTTCAATTAATGAGAAATTAATGAAATCAACAAA 960
QY 901 ACACCATTTTACATTTGAAGCAACATCAACATCACTGTGTAAGAAAACAGTCAAAATGAT 960
DB 901 ACACCATTTTACATTTGAAGCAACATCAACATCACTGTGTAAGAAAACAGTCAAAATGAT 960
QY 954 ACACCATTTTACATTTGAAGCAACATCAACATCACTGTGTAAGAAAACAGTCAAAATGAT 954
DB 954 ACACCATTTTACATTTGAAGCAACATCAACATCACTGTGTAAGAAAACAGTCAAAATGAT 954
QY 961 ACCTTAAAGTTGATTAACCAACCAAGCTTATGATGTAAGATGTGCAATTGGCGAA 1020
DB 961 ACCTTAAAGTTGATTAACCAACCAAGCTTATGATGTAAGATGTGCAATTGGCGAA 1020

421 AAAAGAAACCAAAAGAGGTGTAGTGTCTACGAATATGTGTGGCTTCCAGTT 480
421 AAAAGAAACCAAAAGAGGTGTAGTGTCTACGAATATGTGTGGCTTCCAGTT 480
481 TACGAATGATCAGCAACAGATGGTTCCTATAATAATGAAACAGAAATTAAGCGTT 540
481 TACGAATGATCAGCAACAGATGGTTCCTATAATAATGAAACAGAAATTAAGCGTT 540
541 GTTCATATTTATCTTAAATAATGTGTAGCAATGATGTAGTGTACATGTGAAAAAGTA 600
541 GTTCATATTTATCTTAAATAATGTGTAGCAATGATGTAGTGTACATGTGAAAAAGTA 594
601 GGAACGTGTAATAATGAAAGATTAATGCGCAAGATTTGTTATTTCTAAAGCGAAGGC 660
601 GGAACGTGTAATAATGAAAGATTTAAATGCGCAAGATTTGTTATTTCTAAAGCGAAGGC 654
655 TCACCAAGGCAAGTAAATATATATCCAAAGAGTCAAAAGATGATTAATATCATGACCAAG 714
655 TCACCAAGGCAAGTAAATATATATCCAAAGAGTCAAAAGATGATTAATATCATGACCAAG 714
721 GATTAAGAAACAGCAAAAGCTTTATATCTGGGAAAAAGTTAATGAATTTGGCAAAATGAT 780
721 GATTAAGAAACAGCAAAAGCTTTATATCTGGGAAAAAGTTAATGAATTTGGCAAAATGAT 774
745 GATTAAGAAACAGCAAAAGCTTTATATCTGGGAAAAAGTTAATGAATTTGGCAAAATGAT 774
781 TTCACAGAAAGCAGAAATGAAACGGGAGAAATTAACAGTTAAAAATCTTGAGGTGGTTGG 840
775 TTCACAGAAAGCAGAAATGAAACGGGAGAAATTAACAGTTAAAAATCTTGAGGTGGTTGG 834
841 TATATTTTGAAGAAATGTAAGCTCCAAATTAATGCAAAATTAATGAAATCAACAAA 900
841 TATATTTTGAAGAAATGTAAGCTCCAAATTAATGCAAAATTAATGAAATCAACAAA 894
835 TATATTTTGAAGAAATGTAAGCTCCAAATTAATGCAAAATTAATGAAATCAACAAA 894
901 ACACATTTTACAAATGAAAGCAACAACTCAACCTGTGAAAAAACAATCAAAATGAT 960
901 ACACATTTTACAAATGAAAGCAACAACTCAACCTGTGAAAAAACAATCAAAATGAT 954
895 ACACATTTTACAAATGAAAGCAACAACTCAACCTGTGAAAAAACAATCAAAATGAT 954
961 ACCTCTAAAGTGTATTAACCAACAGCTTGAATGTAAAGATGTGGCAATTTGGCGAA 1020
961 ACCTCTAAAGTGTATTAACCAACAGCTTGAATGTAAAGATGTGGCAATTTGGCGAA 1014
955 ACCTCTAAAGTGTATTAACCAACAGCTTGAATGTAAAGATGTGGCAATTTGGCGAA 1014
1021 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAGACAAAGAGCGAC 1080
1021 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAGACAAAGAGCGAC 1074
1015 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGATTTGCAGACAAAGAGCGAC 1074
1081 GCTATTAATAATAGTCAAAATTCATTTAGTGAATTAACAGATGCGACCTTAATCTTTGAT 1140
1075 GCTATTAATAATAGTCAAAATTCATTTAGTGAATTAACAGATGCGACCTTAATCTTTGAT 1134
1141 AACGTGACTTCTGAGAGATATCTTATGCGTTATATGATGGGATTAACAGTATGCTCCT 1200
1141 AACGTGACTTCTGAGAGATATCTTATGCGTTATATGATGGGATTAACAGTATGCTCCT 1194
1135 AACGTGACTTCTGAGAGATATCTTATGCGTTATATGATGGGATTAACAGTATGCTCCT 1194
1201 GAAATTTATCAAGTGAATGAACAGAAATGCGCTTCACTGTGCGGTTAATCAAGCGTAT 1260
1201 GAAATTTATCAAGTGAATGAACAGAAATGCGCTTCACTGTGCGGTTAATCAAGCGTAT 1254
1195 GAAATTTATCAAGTGAATGAACAGAAATGCGCTTCACTGTGCGGTTAATCAAGCGTAT 1254
1261 ATTCCTAGCTTAACACAGGCGGCACTATAAATCGTTTACTTTATGATTTAAATGAA 1320
1261 ATTCCTAGCTTAACACAGGCGGCACTATAAATCGTTTACTTTATGATTTAAATGAA 1314
1255 ATTCCTAGCTTAACACAGGCGGCACTATAAATCGTTTACTTTATGATTTAAATGAA 1314
1311 AAAAGAGATCTTACGAAAGCTTTTAAATAATGAGCGAATGTTGATACCGTATACCGAC 1380
1311 AAAAGAGATCTTACGAAAGCTTTTAAATAATGAGCGAATGTTGATACCGTATACCGAC 1374
1315 AAAAGAGATCTTACGAAAGCTTTTAAATAATGAGCGAATGTTGATACCGTATACCGAC 1374
1381 GACCAAAACCAACCACTGTTGAAGTTGTAAGGTTGGGAAACGTTTCACTTAAAGTCGAT 1440
1381 GACCAAAACCAACCACTGTTGAAGTTGTAAGGTTGGGAAACGTTTCACTTAAAGTCGAT 1434
1375 GACCAAAACCAACCACTGTTGAAGTTGTAAGGTTGGGAAACGTTTCACTTAAAGTCGAT 1434
1441 GCGCATGTGACAGCAACAGCTTTGGGGAGCTTCTTTGTGCTCGGTATCAAAAC 1500
1441 GCGCATGTGACAGCAACAGCTTTGGGGAGCTTCTTTGTGCTCGGTATCAAAAC 1494
1435 GCGCATGTGACAGCAACAGCTTTGGGGAGCTTCTTTGTGCTCGGTATCAAAAC 1494
1501 AGCGACACAGCAAAATTAATTTGAATAATGATGAAACAAACGAAAGCAACCTTGCGTGAAA 1560

1495 AGCGACACAGCAAAATTTTGAATAATGATGAAACAAAGAAAGCAACTTGGGTGAAA 1554
1561 ACAAAAGCTGAAGCACTACTTTTACAACACAGCGTATGATGATATATCAACAGG 1620
1561 ACAAAAGCTGAAGCACTACTTTTACAACACAGCGTATGATGATATATCAACAGG 1614
1555 ACAAAAGCTGAAGCACTACTTTTACAACACAGCGTATGATGATATATCAACAGG 1614
1621 CTTAAATCGGTACTTATTTATTAAGAAATCTAGTCCGATGATTAATGCTTTGTA 1680
1621 CTTAAATCGGTACTTATTTATTAAGAAATCTAGTCCGATGATTAATGCTTTGTA 1674
1615 CTTAAATCGGTACTTATTTATTAAGAAATCTAGTCCGATGATTAATGCTTTGTA 1674
1681 ACAAAATCGATTTGAATTTGTGCTCAATGAACATATATGACAAACAGAAACCTAGTT 1740
1675 ACAAAATCGATTTGAATTTGTGCTCAATGAACATATATGACAAACAGAAACCTAGTT 1734
1741 TCACCAAGAAAGTAACCAACAAACAGGTAACCTTACCTTCAACAGGTGGCAAGGA 1800
1741 TCACCAAGAAAGTAACCAACAAACAGGTAACCTTACCTTCAACAGGTGGCAAGGA 1794
1735 TCACCAAGAAAGTAACCAACAAACAGGTAACCTTACCTTCAACAGGTGGCAAGGA 1794
1801 ATCTAGTTTACTTATGAAAGTGGCGAGTCTGCTACTTATTAATGAGAGTCTACTTGTCT 1860
1795 ATCTAGTTTACTTATGAAAGTGGCGAGTCTGCTACTTATTAATGAGAGTCTACTTGTCT 1854
1861 AGACGTAGAAAAAGAAATGCT 1881
1855 AGACGTAGAAAAAGAAATGCT 1875

RESULT 11
AAK20109
ID AAK20109 standard; DNA; 1687 BP.
XX
AC AAK20109;
XX
DT 20-APR-1999 (first entry)
XX
DE Enterococcus faecalis EF058 gene fragment.
XX
KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
KW detection; attenuation; antigenic; ss.
XX
OS Enterococcus faecalis.
XX
PN WO980554-A2.
XX
PD 12-NOV-1998.
XX
PF 04-MAY-1998; 98WO-US008959.
XX
PR 06-MAY-1997; 97US-0044031P.
PR 16-MAY-1997; 97US-004655P.
PR 14-NOV-1997; 97US-0066009P.
XX
PA (HDNA-) HUMAN GENOME SCI INC.
PI Kunsch CA, Choi GH, Bailey C, Hromockyj A;
XX WPI: 1999-070095/06.
DR P-PSDB; AAY00119.
XX
XX New isolated Enterococcus faecalis polynucleotides - used to develop
PT products for the detection of Enterococcus and for use in vaccines for
PT prevention or attenuation of Enterococcus infection.
XX
PS Claim 1; Page 134; 301p; English.
XX
XX The present sequence encodes an antigenic polypeptide fragment isolated
CC from Enterococcus faecalis. The present invention describes genes,
CC proteins and antigenic polypeptides isolated from E. faecalis. The
CC proteins can be used in vaccines for preventing or attenuating an
CC infection caused by a member of the Enterococcus genus in an animal. They
CC can also be used for detecting Enterococcus antibodies in a sample. The
CC nucleotide sequences can be used for detecting Enterococcus nucleic

CC acids. Products from the present invention can also be used for screening
CC compounds to identify agonists and antagonists of *E. faecalis* protein
CC activity

XX Sequence 1687 BP; 632 A; 284 C; 360 G; 411 T; 0 U; 0 Other;

Query Match 89.7%; Score 1687; DB 2; Length 1687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 96 AGAAGAAAATGGGAGAGCGCACAGCTCGTGAATTCACAAAAGAAAATGACGAAATTTAC 155
DB 1 AGAAGAAAATGGGAGAGCGCACAGCTCGTGAATTCACAAAAGAAAATGACGAAATTTAC 60
QY 156 AGATCCGCTTATTCAAAATAGCGGAAAAGAAATGACGAAATTTATTAATATCAAGACT 215
DB 61 AGATCCGCTTATTCAAAATAGCGGAAAAGAAATGACGAAATTTATTAATATCAAGACT 120
QY 216 GGCAGATGTCAGCTTAAATTTATTAACGTGACGAAAGAAATTTTACGACGACGCGC 275
DB 121 GGCAGATGTCAGCTTAAATTTATTAACGTGACGAAAGAAATTTTACGACGACGCGC 180
QY 276 AGCGCAAGCGGTGATGACGCTAAACAAGCTGTCCAAAGTTTAACTCTGGGAAACTGT 335
DB 181 AGCGCAAGCGGTGATGACGCTAAACAAGCTGTCCAAAGTTTAACTCTGGGAAACTGT 240
QY 336 TGCTCAAGAAACCAACCGATGCAATGCGAATGTCAGTGTTCAGTTAATCTAAACAAA 395
DB 241 TGCTCAAGAAACCAACCGATGCAATGCGAATGTCAGTGTTCAGTTAATCTAAACAAA 300
QY 396 TGGTAAAGATGACGATGATACCATTAAGAAAGAAACCAAAAGGGGTGATGTCGCTAC 455
DB 301 TGGTAAAGATGACGATGATACCATTAAGAAAGAAACCAAAAGGGGTGATGTCGCTAC 360
QY 456 GAATATGCTGTCGCGTCCAGTTCAGAAATGATCAGCAACAAGATGTTCTTATTA 515
DB 361 GAATATGCTGTCGCGTCCAGTTCAGAAATGATCAGCAACAAGATGTTCTTATTA 420
QY 516 ATATGGAACAGAAAGAAATGCGGTGTCATTTATCTTAAATGTCGTAAGCAATGA 575
DB 421 ATATGGAACAGAAAGAAATGCGGTGTCATTTATCTTAAATGTCGTAAGCAATGA 480
QY 576 TGGTAAATGATGAAAGAAAGTAGAACTGCTGAAGTAAGGATTAATGCGCGACA 635
DB 481 TGGTAAATGATGAAAGAAAGTAGAACTGCTGAAGTAAGGATTAATGCGCGACA 540
QY 636 ATTTGTTATTTCTAAAGCGAAGCTCAACAAGCAAGTAAATATATCAAGAGTCA 695
DB 541 ATTTGTTATTTCTAAAGCGAAGCTCAACAAGCAAGTAAATATATCAAGAGTCA 600
QY 696 AGATGATTAATATACATGACACGATTAAGAAAGACAGCAAAACGCTTTATTA 755
DB 601 AGATGATTAATATACATGACACGATTAAGAAAGACAGCAAAACGCTTTATTA 660
QY 756 AGTTAATGAATGGCGAAAATGATTTCAAGAGCAGAGATGGAAGCGGGAATTAAC 815
DB 661 AGTTAATGAATGGCGAAAATGATTTCAAGAGCAGAGATGGAAGCGGGAATTAAC 720
QY 816 AGTTAATGAATCTTGAAGTTGCTGATATTTTAAAGAAAGTAAAGCTCCAAATATGC 875
DB 721 AGTTAATGAATCTTGAAGTTGCTGATATTTTAAAGAAAGTAAAGCTCCAAATATGC 780
QY 876 AGAATTAATGAATCAAAACAAACCATTTTCAATTTGAAGCAATCAACACCC 935
DB 781 AGAATTAATGAATCAAAACAAACCATTTTCAATTTGAAGCAATCAACACCC 840
QY 936 TGTGAAAAAACAAGTCAAAATGATACCTTAAGTTGATTAACCAACCAAGCTTGA 995
DB 841 TGTGAAAAAACAAGTCAAAATGATACCTTAAGTTGATTAACCAACCAAGCTTGA 900
QY 996 TGGTAAAGATGTCGCAATGGCGAAAATTAATATCAAAATTTCTGTAATATTCATT 1055
DB 901 TGGTAAAGATGTCGCAATGGCGAAAATTAATATCAAAATTTCTGTAATATTCATT 960

```

```

QY 1056 GGGATTCGACGCAAGAGAGCGACTTAATATACGTAATCAATTTAGTGAATA 1115
DB 961 GGGATTCGACGCAAGAGAGCGACTTAATATACGTAATCAATTTAGTGAATA 1020
QY 1116 ACATGATCAGCGCTTAATCTTTTGAATACGTACTTGGAGATGATGCTTATATG 1175
DB 1021 ACATGATCAGCGCTTAATCTTTTGAATACGTACTTGGAGATGATGCTTATATG 1080
QY 1176 TGATGGGATACAGTGAATGCTCTGTAATTAATATCAAGTGAACAGCAATGCTT 1235
DB 1081 TGATGGGATACAGTGAATGCTCTGTAATTAATATCAAGTGAACAGCAATGCTT 1140
QY 1236 CACTGTCGCGCTTAATCCAGGCTATATTCCTACGCTAACCAAGCGGCACTAAAT 1295
DB 1141 CACTGTCGCGCTTAATCCAGGCTATATTCCTACGCTAACCAAGCGGCACTAAAT 1200
QY 1296 CGTTTACTTTATGACATTTAAATGAAGAAAGCAGATCCTACGAAGGCTTTAAATGAGGC 1355
DB 1201 CGTTTACTTTATGACATTTAAATGAAGAAAGCAGATCCTACGAAGGCTTTAAATGAGGC 1260
QY 1356 GAATGTTGATTAACGTCATACCGACGACCAACCAACCAACTGTTGAAGTTGACAG 1415
DB 1261 GAATGTTGATTAACGTCATACCGACGACCAACCAACCAACTGTTGAAGTTGACAG 1320
QY 1416 TGGAAACGTTTCATTAAGTCGATGCGGATGTGACACGACACAAAGCTTGGCGGAGC 1475
DB 1321 TGGAAACGTTTCATTAAGTCGATGCGGATGTGACACGACACAAAGCTTGGCGGAGC 1380
QY 1476 TTCCTTGTGCTGCGTGAATCAAAACAGGCAACAGCAATTAATTTGAATGATGAAC 1535
DB 1381 TTCCTTGTGCTGCGTGAATCAAAACAGGCAACAGCAATTAATTTGAATGATGAAC 1440
QY 1536 AACGAAACGACCAACTGGTGAAGAAACAAAGCTGAACCACTTTTACAAACGCGC 1595
DB 1441 AACGAAACGACCAACTGGTGAAGAAACAAAGCTGAACCACTTTTACAAACGCGC 1500
QY 1596 TGATGATTAATGATATCAACAGGCTTAATATGCGTACTTATTTGAAGAAAGCTGT 1655
DB 1501 TGATGATTAATGATATCAACAGGCTTAATATGCGTACTTATTTGAAGAAAGCTGT 1560
QY 1656 AGCTCCGATGATTAATGCTTGTTPAACAAATCGAATTAATTTGAGTCAATGAACAATC 1715
DB 1561 AGCTCCGATGATTAATGCTTGTTPAACAAATCGAATTAATTTGAGTCAATGAACAATC 1620
QY 1716 ATATGCGACAAACAGAAAACCTAGTTTCAACGAAAAGTATACCAACCAAGGTAC 1775
DB 1621 ATATGCGACAAACAGAAAACCTAGTTTCAACGAAAAGTATACCAACCAAGGTAC 1680
QY 1776 CTTAAGCT 1782
DB 1681 CTTAAGCT 1687

```

RESULT 12

ID ABN98094 standard; DNA; 1687 BP.

XX ABN98094;

DT 05-AUG-2002 (first entry)

DE *E. faecalis* EF058 gene fragment.

XX Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic;

XX gene; *del.*

XX Enterococcus faecalis.

FN US2002045737-A1.

PD 18-APR-2002.

PF 04-MAY-1998; 98US-00071035.
XX 04-MAY-1998; 98US-00071035.
XX (HUMA-) HUMAN GENOME SCI INC.
PA ChOI GH, Bailey C, Hromocky J A, Kunsch CA;
XX WPI, 2002-425450/45.
DR P-PSDB; ABP4338.
XX
XX New genes and polypeptides from Enterococcus faecalis, useful as vaccines
PT for preventing, treating or attenuating an infection caused by a member
PT of the Enterococcus genus in an animal, particularly E. faecalis.
XX
XX Claim 1; Page 93-94; 255pp; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of polypeptides from Enterococcus faecalis. The proteins can be
CC used as vaccines for preventing or attenuating an infection caused by a
CC member of the Enterococcus genus in an animal, particularly E. faecalis.
CC The polynucleotide is also useful for preventing or treating E. faecalis
CC infection. The present sequence is a coding sequence of the invention
XX
SQ Sequence 1687 BP; 632 A; 284 C; 360 G; 411 T; 0 U; 0 Other;
Query Match 89.7%; Score 1687; DB 6; Length 1687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 96 AGAAGAAATGCGGAGAGCGACAGCTCGTATTCACAAAAGAAAATGACGGATTTCAC 155
DB 1 AGAAGAAATGCGGAGAGCGACAGCTCGTATTCACAAAAGAAAATGACGGATTTCAC 60
QY 156 AGATCCGCTTATTCAAAATAGCGGAGAAAATGACGGATTTCACAAAATGACGGACT 215
DB 61 AGATCCGCTTATTCAAAATAGCGGAGAAAATGACGGATTTCACAAAATGACGGACT 120
QY 216 GGCAGATGTCAGCTTATTTATTAAGTCGACGAAATTTTCGACGACGACGCGC 275
DB 121 GGCAGATGTCAGCTTATTTATTAAGTCGACGAAATTTTCGACGACGACGCGC 180
QY 276 AGGCGCAAGCGTTGTCAGCTTAAACAAGCTGTCAAAGTTTACTCTGGGAAACTGT 335
DB 181 AGGCGCAAGCGTTGTCAGCTTAAACAAGCTGTCAAAGTTTACTCTGGGAAACTGT 240
QY 336 TGCTCAGGAACCCAGATGCAATGGGAATGTCACTGTTCACTTAACCTTAAACAAA 395
DB 241 TGCTCAGGAACCCAGATGCAATGGGAATGTCACTGTTCACTTAACCTTAAACAAA 300
QY 396 TGGTAAAGATGACGATTAACATTAAGAAAGAACCAAAAGGGGTGATGTCGCTAC 455
DB 301 TGGTAAAGATGACGATTAACATTAAGAAAGAACCAAAAGGGGTGATGTCGCTAC 360
QY 456 GAATATGTCGTCGCTTCCAGTTTACGAAATGATCAAGCAACAGATGTCCTATA 515
DB 361 GAATATGTCGTCGCTTCCAGTTTACGAAATGATCAAGCAACAGATGTCCTATA 420
QY 516 ATATGGAACAGAAATAGCGGTGTTCTATTTATCTTAAATATGTCGCAATGA 575
DB 421 ATATGGAACAGAAATAGCGGTGTTCTATTTATCTTAAATATGTCGCAATGA 480
QY 576 TGGTATGTTACATGTAAGAAAGTAGAAGCTGTAAGATGAAAGTTTAAATGCGCGCA 635
DB 481 TGGTATGTTACATGTAAGAAAGTAGAAGCTGTAAGATGAAAGTTTAAATGCGCGCA 540
QY 636 ATTTGTTATTTCTTAAAGGAGCTCAACGACCAAGTAAATATATCAAGAGTCA 695
DB 541 ATTTGTTATTTCTTAAAGGAGCTCAACGACCAAGTAAATATATCAAGAGTCA 600
QY 696 AGATGATTTATATCAATGACCAAGATTAAGAAACAAAGCAACGTTTATTAATCTGGAA 755
DB 601 AGATGATTTATATCAATGACCAAGATTAAGAAACAAAGCAACGTTTATTAATCTGGAA 660

QY 756 AAGTTAAGAAATGCGGAGAAATGATTTCAAGAACAGAGATGGAACGGGAAATTAAC 815
DB 661 AAGTTAAGAAATGCGGAGAAATGATTTCAAGAACAGAGATGGAACGGGAAATTAAC 720
QY 816 AGTTAAATATCTTGAGTTGGTTCGTATATTTAGAAAGTAAAGTCCAAATATGC 875
DB 721 AGTTAAATATCTTGAGTTGGTTCGTATATTTAGAAAGTAAAGTCCAAATATGC 780
QY 876 AGAATTAATGAAATCAAAACAAACCCATTTACATTTGAAGCAACATCAACACC 935
DB 781 AGAATTAATGAAATCAAAACAAACCCATTTACATTTGAAGCAACATCAACACC 840
QY 936 TGTGAAAAAACAAGCAAAATGATACCTTAAAGTTGATTAACCAACCAAGCTTGA 995
DB 841 TGTGAAAAAACAAGCAAAATGATACCTTAAAGTTGATTAACCAACCAAGCTTGA 900
QY 996 TGTAAAGATGTCGCAATGCGGAGAAATTAATATCAAAATTTCTGTAAATTCAT 1055
DB 901 TGTAAAGATGTCGCAATGCGGAGAAATTAATATCAAAATTTCTGTAAATTCAT 960
QY 1056 GGGATTTGACCAAGAAAGGCGACGCTTAATTAATACCTCAATTTATAGTTGATTA 1115
DB 961 GGGATTTGACCAAGAAAGGCGACGCTTAATTAATACCTCAATTTATAGTTGATTA 1020
QY 1116 ACATGATGACGCTTAACTTTGATTAAGTACCTTGAAGATGCTTAATGCTTATA 1175
DB 1021 ACATGATGACGCTTAACTTTGATTAAGTACCTTGAAGATGCTTATGCTTATA 1080
QY 1176 TGATGGGATACAGTATGCTCTGAAATATTAATCAATGACGAAACGAATGCGTT 1235
DB 1081 TGATGGGATACAGTATGCTCTGAAATATTAATCAATGACGAAACGAATGCGTT 1140
QY 1236 CACTGCGCGTTTATTCAGGATTAATTCCTACGTAACACAGGCGGACCTAAAT 1295
DB 1141 CACTGCGCGTTTATTCAGGATTAATTCCTACGTAACACAGGCGGACCTAAAT 1200
QY 1296 CGTTTACTTATGCAATTAATGAAAGCAAGATCTACGAAGGCTTTAAATATGAGC 1355
DB 1201 CGTTTACTTATGCAATTAATGAAAGCAAGATCTACGAAGGCTTTAAATATGAGC 1260
QY 1356 GAATGTTGATACCGTCAATCCAGACCAACACCAACCTGTTGAATGTGACAG 1415
DB 1261 GAATGTTGATACCGTCAATCCAGACCAACACCAACCTGTTGAATGTGACAG 1320
QY 1416 TGGGAAGCTTCAATTAAGTCGATGGGATGTCACGACACAAAGCTTGGCGGAGC 1475
DB 1321 TGGGAAGCTTCAATTAAGTCGATGGGATGTCACGACACAAAGCTTGGCGGAGC 1380
QY 1476 TTCTTTGTCGTCGATCAAAACAGCGACACAGCAATTAATTAATGAAATGATGAAC 1535
DB 1381 TTCTTTGTCGTCGATCAAAACAGCGACACAGCAATTAATTAATGAAATGATGAAC 1440
QY 1536 AACGAAGCAGCACTTGGGTGAAAAAAGAAAGTGAAGCACTTTTCAACAGCGC 1595
DB 1441 AACGAAGCAGCACTTGGGTGAAAAAAGAAAGTGAAGCACTTTTCAACAGCGC 1500
QY 1596 TGAATGATGATGTAATCAACAGGCTTAAATAGGCTTAAATTTAGAAAGAACTGT 1655
DB 1501 TGAATGATGATGTAATCAACAGGCTTAAATAGGCTTAAATTTAGAAAGAACTGT 1560
QY 1656 AGCTCTGATGATTAATGCTTGTAAACAAATCGAATGAAATTTGTCATGAACATC 1715
DB 1561 AGCTCTGATGATTAATGCTTGTAAACAAATCGAATGAAATTTGTCATGAACATC 1620
QY 1716 ATATGCAACACAGAAACCTTATGTTTACCGAAAGAAAGTACCAACCAAGGTC 1775
DB 1621 ATATGCAACACAGAAACCTTATGTTTACCGAAAGAAAGTACCAACCAAGGTC 1680
QY 1776 CTAACT 1782
DB 1681 CTAACT 1687

RESULT 13
ACA88058
ID ACA88058 standard; DNA; 1687 BP.
XX
AC ACA88058;
XX
DT 07-JUL-2003 (first entry)
XX
DE E. faecalis novel gene #110.
XX
KM Gene; ds; endocarditis; bacteraemia; urinary tract infection; URI;
KM intrabdominal infection; soft tissue infection; neonatal sepsis;
KM vaccine.
XX
OS Enterococcus faecalis.
XX
PN US2003017495-A1.
XX
PD 23-JAN-2003.
XX
PF 29-JUL-2002; 2002US-00206576.
XX
PR 06-MAY-1997; 97US-0044031P.
PR 16-MAY-1997; 97US-0046655P.
PR 14-NOV-1997; 97US-0066009P.
PR 04-MAY-1998; 98US-00071035.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;
XX
DR WPI; 2003-416890/39.
DR P-PsDB; ADB88366.
XX
PT New nucleic acid molecules and polypeptides from Enterococcus faecalis,
PT useful as vaccines for preventing or attenuating an enterococcal
PT infection in an animal, or for identifying Enterococcus faecalis in
PT biological samples.
XX
PS Claim 1; Page; 40pp; English.
XX
CC The invention relates to a new isolated nucleic acid molecule comprising
CC a polynucleotide isolated from Enterococcus faecalis appearing as
CC ACA87949-ACA88196 (or sequences complementary to them or 95% identical to
CC them). Also included are the proteins encoded by the above nucleic acids,
CC making a recombinant vector (comprising inserting the isolated nucleic
CC acid molecule cited above into a vector), a host cell comprising the
CC vector (used to produce the protein), an isolated antibody specific for
CC the polypeptides, a hybridoma that produces the antibody, an isolated
CC polypeptide antigen comprising an amino acid sequence of an Enterococcus
CC faecalis epitope listed in the specification, a vaccine comprising one or
CC more E. faecalis polypeptides (and a pharmaceutical diluent, carrier or
CC excipient) where the polypeptide elicits protective antibodies in an
CC animal to a member of the genus Enterococcus; preventing or attenuating
CC an infection caused by a member of the genus Enterococcus in an animal
CC comprising administering to the animal the polypeptide and detecting
CC Enterococcus nucleic acids in a biological sample. The E. faecalis
CC nucleic acid molecules and polypeptides are useful as vaccines for
CC preventing or attenuating an enterococcal infection in an animal (e.g.
CC endocarditis, bacteraemia, urinary tract infection (UTI), intrabdominal
CC infection, soft tissue infection and neonatal sepsis). The polypeptides
CC are also useful for detecting Enterococcus aureus in immunosays, as
CC epitope tags, as molecular weight markers, or for generating antibodies
CC that specifically bind E. faecalis polypeptides. The nucleic acid
CC molecules are also useful as probes for gene mapping, or for identifying
CC E. faecalis in biological samples. The kit and methods are useful for
CC detecting Enterococcus antibodies or nucleic acid molecules in a
CC biological sample. The present sequence is a novel E. faecalis nucleic
CC acid of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030017495

XX
SQ Sequence 1687 BP; 632 A; 284 C; 360 G; 411 T; 0 U; 0 Other;
Query Match 89.7%; Score 1687; DB 8; Length 1687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY
QY 96 AGAAGAAAATGGGAGAGGCGACAGCTGTGATTCACAAAAAGAAATACGGAATTAC 155
DB 1 AGAAGAAAATGGGAGAGGCGACAGCTGTGATTCACAAAAAGAAATACGGAATTAC 60
QY 156 AGATCGCTTATTCACAAATAGCGGAGAAATAGCGAGTTGATTAATACAGACT 215
DB 61 AGATCGCTTATTCACAAATAGCGGAGAAATAGCGAGTTGATTAATACAGACT 120
QY 216 GGCAGATGTGACGTTAGTATTATTAACGTGACGAAAGCAATTTTACGACACGAGCGC 275
DB 121 GGCAGATGTGACGTTAGTATTATTAACGTGACGAAAGCAATTTTACGACACGAGCGC 180
QY 276 AGGCGCAAGCGTTGATGACGCTTAACAGCTGTCCAAAGTTTAACCTCTGGGAAACCTGT 335
DB 181 AGGCGCAAGCGTTGATGACGCTTAACAGCTGTCCAAAGTTTAACCTCTGGGAAACCTGT 240
QY 336 TGCTCAAGGAACCAACGATGCAATGCGAATGTCTGCTTCACTTAACCAAAAA 395
DB 241 TGCTCAAGGAACCAACGATGCAATGCGAATGTCTGCTTCACTTAACCAAAAA 300
QY 396 TGGTAAGATGCAAGTGTATACCATTAAGAAAGCAAAAGAGGAGTGTAGTGTCTGCTAC 455
DB 301 TGGTAAGATGCAAGTGTATACCATTAAGAAAGCAAAAGAGGAGTGTAGTGTCTGCTAC 360
QY 456 GAATATGTGTGTGGGCTTCCAGTTTACGAAATGATCAAGCAAAAGAGTGTCTGCTAC 515
DB 361 GAATATGTGTGTGGGCTTCCAGTTTACGAAATGATCAAGCAAAAGAGTGTCTGCTAC 420
QY 516 ATATGAAACAGAAAGATAGCGGTTTCATTTATCTTAATAATGTGTGCTCAATGA 575
DB 421 ATATGAAACAGAAAGATAGCGGTTTCATTTATCTTAATAATGTGTGCTCAATGA 480
QY 576 TGGTAGTTTACATGTGAAAAAAGTAGAAGCTGTGAAATGAGGATTAATGGCGCAGA 635
DB 481 TGGTAGTTTACATGTGAAAAAAGTAGAAGCTGTGAAATGAGGATTAATGGCGCAGA 540
QY 636 ATTTGTTATTTCTTAACCGAAGCTGACGAGCAATTAATATTCAGAGAGTCA 695
DB 541 ATTTGTTATTTCTTAACCGAAGCTGACGAGCAATTAATATTCAGAGAGTCA 600
QY 696 AGATGATTTATATACATGAGCAACGATTAAGAAACAGCAAAAAAGCTTATTAAGTGGGA 755
DB 601 AGATGATTTATATACATGAGCAACGATTAAGAAACAGCAAAAAAGCTTATTAAGTGGGA 660
QY 756 AAGTTATGAAATTTGGCGAAAAATGATTTCACAGAACGAGATGGAACGGAGAAATTAC 815
DB 661 AAGTTATGAAATTTGGCGAAAAATGATTTCACAGAACGAGATGGAACGGAGAAATTAC 720
QY 816 AGTTAAATATCTGAGGTGGTTCGTATATTTTAAAGAAAGTAAAGCTCCAAATTAATGC 875
DB 721 AGTTAAATATCTGAGGTGGTTCGTATATTTTAAAGAAAGTAAAGCTCCAAATTAATGC 780
QY 876 AGAATTAATTAATCAAAACAAACAAACAAACCAATTCATTTGAGCAATCAACAC 935
DB 781 AGAATTAATTAATCAAAACAAACAAACCAATTCATTTGAGCAATCAACAC 840
QY 936 TGGTGAAGAAACAGTCAAAATGATACCTTAAGTTGATTAACACACCAAGCTTGA 995
DB 841 TGGTGAAGAAACAGTCAAAATGATACCTTAAGTTGATTAACACACCAAGCTTGA 900
QY 996 TGGTGAAGAAATGGCAATTTGGCGAAATTAATTAATTAATTAATTAATTAATTCCTT 1055
DB 901 TGGTGAAGAAATGGCAATTTGGCGAAATTAATTAATTAATTAATTAATTCCTT 960
QY 1056 GGGGATTCAGACAAAGAAAGCGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1115


```

Db      541 ATTGTTATTTCTAAAAAGGAGGCTCACAGGACAGATATAATATCAAGAGTCAA 600
Qy      696 AGATGATTTATATCATGAGCAACGATTAAGAAACAAGCAAAACGTTTATTACTGGGAA 755
Db      601 AGATGATTTATATCATGAGCAACGATTAAGAAACAAGCAAAACGTTTATTACTGGGAA 660
Qy      756 AAGTTATGAAATTTGGGAAATGATTTTCAGAGAGAGAAATGAAACGGAGAAATTAAC 815
Db      661 AAGTTATGAAATTTGGGAAATGATTTTCAGAGAGAGAAATGAAACGGAGAAATTAAC 720
Qy      816 AGTTAAAAATCTTAGAGTTGTTCTGTATATTTAGAGAGATTAAGCTCCAAATATGC 875
Db      721 AGTTAAAAATCTTAGAGTTGTTCTGTATATTTAGAGAGATTAAGCTCCAAATATGC 780
Qy      876 AGAATTTAATTTGAAAAATCAACAAAAACACATTTTACATTTGAAGCAACATCAACAC 935
Db      781 AGAATTTAATTTGAAAAATCAACAAAAACACATTTTACATTTGAAGCAACATCAACAC 840
Qy      936 TGTGAAAAAAGTCATAAATGATACCTCTAAAGTTGATTAACAAACCAAGCTTAGA 995
Db      841 TGTGAAAAAAGTCATAAATGATACCTCTAAAGTTGATTAACAAACCAAGCTTAGA 900
Qy      996 TGTGAAAGATGTGGCAATTTGGGAAAAAATTAATATCAATTTCTGTAAATATTCAT 1055
Db      901 TGTGAAAGATGTGGCAATTTGGGAAAAAATTAATATCAATTTCTGTAAATATTCAT 960
Qy      1056 GGGGATTTGAGCAAAAGAGGCGCTAATATATCGTCAATTTCAATTTAGTTGATTA 1115
Db      961 GGGGATTTGAGCAAAAGAGGCGCTAATATATCGTCAATTTCAATTTAGTTGATTA 1020
Qy      1116 ACATGATGAGGCTTAACTTTGATATACGTGACTTGTGAGAGTATGCTTAAGCTTAA 1175
Db      1021 ACATGATGAGGCTTAACTTTGATATACGTGACTTGTGAGAGTATGCTTAAGCTTAA 1080
Qy      1176 TGATGGGATACAGTGAATGCTCTGAAAAATTAATCAAGTGAACAAAGCAATGGCT 1235
Db      1081 TGATGGGATACAGTGAATGCTCTGAAAAATTAATCAAGTGAACAAAGCAATGGCT 1140
Qy      1236 CACTGTCCGCGTTAATCCAGGCTATATTCCTACGCTAACCAAGGCGGACACTTAAAT 1295
Db      1141 CACTGTCCGCGTTAATCCAGGCTATATTCCTACGCTAACCAAGGCGGACACTTAAAT 1200
Qy      1296 CGTTACTTTATGATTTTAAATGAAAAAGAGATCTCTACGAAAGGCTTAAATATGAGG 1355
Db      1201 CGTTACTTTATGATTTTAAATGAAAAAGAGATCTCTACGAAAGGCTTAAATATGAGG 1260
Qy      1356 GAATGTGATTAACGGTCAATACCGAGCAACCAACCACTGTTGAAGTTGTGACAG 1415
Db      1261 GAATGTGATTAACGGTCAATACCGAGCAACCAACCACTGTTGAAGTTGTGACAG 1320
Qy      1416 TGGGAAACGTTTATTAAGTGTGATGAGGCAATGTCAGGCAACAGCTTTGGCGGAGC 1475
Db      1321 TGGGAAACGTTTATTAAGTGTGATGAGGCAATGTCAGGCAACAGCTTTGGCGGAGC 1380
Qy      1476 TTCCTTTGCTCCGCGTCAAAACAGGCAACAGCAATTTTGAATAATGATGAAC 1535
Db      1381 TTCCTTTGCTCCGCGTCAAAACAGGCAACAGCAATTTTGAATAATGATGAAC 1440
Qy      1536 AACGAAGCAGCACTTGGGTGAAAAACAAGCTGAGCACTTTTACCAACAGCGC 1595
Db      1441 AACGAAGCAGCACTTGGGTGAAAAACAAGCTGAGCACTTTTACCAACAGCGC 1500
Qy      1596 TGAATGATTAAGTTATATCAACGGGCTTAAATACGGTACCTATTTATTTAGAAAGTGT 1655
Db      1501 TGAATGATTAAGTTATATCAACGGGCTTAAATACGGTACCTATTTATTTAGAAAGTGT 1560
Qy      1656 AGCTCTGATGATTAATGCTTTGTTAAACAAATCGAATTTGAGTCAATGAAACATC 1715
Db      1561 AGCTCTGATGATTAATGCTTTGTTAAACAAATCGAATTTGAGTCAATGAAACATC 1620
Qy      1716 ATATGGCAACAGCAAAACCTAGTTTACCAAGAAAAAGTACCAACCAAGAAAGTAC 1775
Db      1621 ATATGGCAACAGCAAAACCTAGTTTACCAAGAAAAAGTACCAACCAAGAAAGTAC 1680

```

```

Qy      1776 CTTACTT 1782
Db      1681 CTTACTT 1687

RESULT 15
ADY38953.
ID      ADY38953 standard; DNA; 1687 BP.
XX
AC      ADY38953;
XX
AC      ADY38953;
XX
DT      05-MAY-2005 (first entry)
XX
DE      Novel Enterococcus faecalis gene sequence SeqID219.
XX
KW      protein purification; DNA purification; antibacterial; vaccine;
KW      enterococcus faecalis infection; gene; de.
XX
OS      Enterococcus faecalis.
XX
PN      US2005043528-A1.
XX
PD      24-FEB-2005.
XX
PF      06-AUG-2004; 2004US-00912362.
XX
PR      06-MAY-1997; 97US-0044031P.
PR      16-MAY-1997; 97US-0046555P.
PR      14-NOV-1997; 97US-0066009P.
PR      04-MAY-1998; 98US-00071035.
PR      29-JUL-2002; 2002US-00206576.
XX
PA      (HUMA-) HUMAN GENOME SCI INC.
XX
PI      Choi GH, Bailey C, Hromocky J A, Kunsch CA;
XX
DR      MPI: 2005-180840/19.
XX
DR      P-PSDB; ADY38954.
XX
PT      New isolated Enterococcus faecalis nucleic acid useful for diagnosing,
PT      preventing and/or attenuating infection by Enterococcus faecalis in
PT      animals, particularly in humans.
XX
PS      Claim 1; SEQ ID NO 219; 28bp; English.
XX
CC      This invention relates to novel genes and the proteins they encode
CC      isolated from Enterococcus faecalis. The invention may be useful for the
CC      development of compounds with an antibacterial activity or a vaccine. The
CC      present invention is useful for the diagnosis, prevention and/or
CC      attenuation of infection by Enterococcus faecalis in animals,
CC      particularly in humans. The present sequence is that of a novel
CC      Enterococcus faecalis gene of the invention. Note: The present sequence
CC      does not form part of the printed specification but was obtained in
CC      electronic form from the USPTO web site.
XX
SQ      Sequence 1687 BP; 632 A; 284 C; 360 G; 411 T; 0 U; 0 Other;

Query Match      89.7%; Score 1687; DB 14; Length 1687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      96 AGAAGAAAAATGGGAGAGAGCGCACAGCTGTGATTTCAAAAAAGAAATGACGGAATTAAC 155
Db      1 AGAAGAAAAATGGGAGAGAGCGCACAGCTGTGATTTCAAAAAAGAAATGACGGAATTAAC 60
Qy      156 AGATCCGCTTATTTCAAAATTAAGCGGAGAAAGAAATGAGCAAGTTGATTAATCAAGACT 215
Db      61 AGATCCGCTTATTTCAAAATTAAGCGGAGAAAGAAATGAGCAAGTTGATTAATCAAGACT 120
Qy      216 GGCAGATGTGACGTTTATTTATTAATACGTGACGAACGAATTTTACAGAGCAAGCGGC 275
Db      121 GGCAGATGTGACGTTTATTTATTAATTAATACGTGACGAACGAATTTTACAGAGCAAGCGGC 180

```

QY 276 AGGCCAAGCGTTGATGCACTTAACAAGCTGTCCAAAGTTAACTCTCGGAAAACCTGT 335
 Db 181 AGGGCAAGCGTTGATGCACTTAACAAGCTGTCCAAAGTTAACTCTCGGAAAACCTGT 240
 QY 336 TGCCTAAGGAACCAACCGATGCAATGGGAATGTCACTGTTCACTTACCTAAACAAACAA 395
 Db 241 TGCCTAAGGAACCAACCGATGCAATGGGAATGTCACTGTTCACTTACCTAAACAAACAA 300
 QY 396 TGGTAAAGATGCACTGTATACCTATTAAAGAAACCAAAAGGGGTGATGTTGCTGTAC 455
 Db 301 TGGTAAAGATGCACTGTATACCTATTAAAGAAACCAAAAGGGGTGATGTTGCTGTAC 360
 QY 456 GAATGTGTTGGTGGCTTCCAGCTTTAGAAATGATCAAGCAAAAGATGTTCTATTA 515
 Db 361 GAATGTGTTGGTGGCTTCCAGCTTTAGAAATGATCAAGCAAAAGATGTTCTATTA 420
 QY 516 ATATGGAACAGAAAGATTAGCGGTGTTCAATTTATCTTAAAAAATGTGTAGCCAAATGA 575
 Db 421 ATATGGAACAGAAAGATTAGCGGTGTTCAATTTATCTTAAAAAATGTGTAGCCAAATGA 480
 QY 576 TGTGATGTTACATGTGAATAAAAGTGAAGCTGCTGAAATGAAAGATTAAATGGCGCAGA 635
 Db 481 TGTGATGTTACATGTGAATAAAAGTGAAGCTGCTGAAATGAAAGATTAAATGGCGCAGA 540
 QY 636 ATTTGTTATTTCTTAAAGCGAAGGCTCACAGGCAAGTAAATATATCAAGAGTCA 695
 Db 541 ATTTGTTATTTCTTAAAGCGAAGGCTCACAGGCAAGTAAATATATCAAGAGTCA 600
 QY 696 AGATGATTTATATACATGACAAACGGATTAAGAACCAAGCAAAACGCTTTATTTACTGGGA 755
 Db 601 AGATGATTTATATACATGACAAACGGATTAAGAACCAAGCAAAACGCTTTATTTACTGGGA 660
 QY 756 AAGTTATGAATTTGGCGAAAAATGATTTCCAGAACAGAGATGGAACGGGAGATTAC 815
 Db 661 AAGTTATGAATTTGGCGAAAAATGATTTCCAGAACAGAGATGGAACGGGAGATTAC 720
 QY 816 AGTTAAATCTTGAAGTTGGTTCGATTTTGAAGAAAGTAAAGTCCAAATATATGC 875
 Db 721 AGTTAAATCTTGAAGTTGGTTCGATTTTGAAGAAAGTAAAGTCCAAATATATGC 780
 QY 876 AGAATTTAATTTGAAAAATCAAAACAAAAACACATTTTCAATTTGAAGCAAAACATCAACACC 935
 Db 781 AGAATTTAATTTGAAAAATCAAAACAAAAACACATTTTCAATTTGAAGCAAAACATCAACACC 840
 QY 936 TGTGAAAAAAACAGTCAAAAAATGATACCTCTTAAAGTTGATTAACAAACCAAGCTTGA 995
 Db 841 TGTGAAAAAAACAGTCAAAAAATGATACCTCTTAAAGTTGATTAACAAACCAAGCTTGA 900
 QY 996 TGGTAAAGATGTCGCAATTTGGGAAAAAATTAATATCAATTTTCTGTAAATATTTCCATT 1055
 Db 901 TGGTAAAGATGTCGCAATTTGGGAAAAAATTAATATCAATTTTCTGTAAATATTTCCATT 960
 QY 1056 GGGGAATTGACAAAGAAAGGCGCTAATATAATACGTCAAATTTCAATTTTATGTTGATTA 1115
 Db 961 GGGGAATTGACAAAGAAAGGCGCTAATATAATACGTCAAATTTCAATTTTATGTTGATTA 1020
 QY 1116 ACATGATGACAGCTTAACTTTTGAATTAAGTGACTTCTGAGAGATAGCTTAAGCGTTATA 1175
 Db 1021 ACATGATGACAGCTTAACTTTTGAATTAAGTGACTTCTGAGAGATAGCTTAAGCGTTATA 1080
 QY 1176 TGAATGGGATACAGTGAATGCTCTGAAAAATATCAAGTGAAGTGAACAAAGCAAAATGGCTT 1235
 Db 1081 TGAATGGGATACAGTGAATGCTCTGAAAAATATCAAGTGAAGTGAACAAAGCAAAATGGCTT 1140
 QY 1236 CACTGTGCGCTTAATCAAGCTATATTCCTAAGCTAAACCAAGCGGCAACTAAATTT 1295
 Db 1141 CACTGTGCGCTTAATCAAGCTATATTCCTAAGCTAAACCAAGCGGCAACTAAATTT 1200
 QY 1296 CGTTTAACTTTATGCAATTTAAAGAAACAGAGATCTTACGAAAGGCTTTAAATGAGGCG 1355
 Db 1201 CGTTTAACTTTATGCAATTTAAAGAAACAGAGATCTTACGAAAGGCTTTAAATGAGGCG 1260

QY 1356 GAATGTTGATACGGTCAATACCGACAGCAACAAACCAACTGTTGAAGTTGTGACAGG 1415
 Db 1261 GAATGTTGATACGGTCAATACCGACAGCAACAAACCAACTGTTGAAGTTGTGACAGG 1320
 QY 1416 TGGGAACGTTTCATTTAAAGTCGATGCGGATGTGACAGCAACAAAGCTTGGCGGAGC 1475
 Db 1321 TGGGAACGTTTCATTTAAAGTCGATGCGGATGTGACAGCAACAAAGCTTGGCGGAGC 1380
 QY 1476 TTCCTTTGCTCCGTCGATCAAAAACAGGCAACAGCAAAATTTATTTGAATAATGATGAAG 1535
 Db 1381 TTCCTTTGCTCCGTCGATCAAAAACAGGCAACAGCAAAATTTATTTGAATAATGATGAAG 1440
 QY 1536 AACGAACACAACTTGGGTGAACCAAAAGCTGAACCACTACTTTTAAACAAACGCG 1595
 Db 1441 AACGAACACAACTTGGGTGAACCAAAAGCTGAACCACTACTTTTAAACAAACGCG 1500
 QY 1596 TGATGATTTAGTTGATATCAAGGCTTTAAATACGGTACCTATTTATTTGAAGAAACCTGT 1655
 Db 1501 TGATGATTTAGTTGATATCAAGGCTTTAAATACGGTACCTATTTATTTGAAGAAACCTGT 1560
 QY 1656 AGCTTCGATGATTTATGCTTTGTTAAACAAATCGATTTGAATTTGTCAATGAACAATC 1715
 Db 1561 AGCTTCGATGATTTATGCTTTGTTAAACAAATCGATTTGAATTTGTCAATGAACAATC 1620
 QY 1716 ATATGCAACAAAGAAAACCTAGTTTCAACAGAAAAAGTACCAACAAACAAAGGTAC 1775
 Db 1621 ATATGCAACAAAGAAAACCTAGTTTCAACAGAAAAAGTACCAACAAACAAAGGTAC 1680
 QY 1776 CTTTACT 1782
 Db 1681 CTTTACT 1687

Search completed: July 5, 2006, 18:22:45
 Job time : 832 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	1881	100.0	1881	8	US-10-661-809-12	Sequence 12, Appl
2	1881	100.0	1983	3	US-09-071-035-217	Sequence 217, Appl
3	1881	100.0	1983	6	US-10-206-576-217	Sequence 217, Appl
4	1881	100.0	1983	9	US-10-912-362-217	Sequence 217, Appl
5	1881	100.0	15614	3	US-09-071-0272A-45	Sequence 45, Appl
6	1687	89.7	1687	3	US-09-071-035-219	Sequence 219, Appl
7	1687	89.7	1687	6	US-10-206-576-219	Sequence 219, Appl
8	1687	89.7	1687	9	US-10-912-362-219	Sequence 219, Appl
9	1014.6	53.9	1878	8	US-10-661-809-18	Sequence 18, Appl
10	595	31.6	2199	8	US-10-333-002-28	Sequence 28, Appl
11	67	3.6	1062	9	US-10-425-115-120013	Sequence 120013, Appl
12	65	3.5	627	8	US-10-931-323-9336	Sequence 9336, Appl
13	64.6	3.4	6609	7	US-10-172-502-1	Sequence 1, Appl
14	64.6	3.4	6609	13	US-11-020-509-1	Sequence 1, Appl
15	64.4	3.4	1029	9	US-10-445-115-123086	Sequence 123086, Appl
16	64.2	3.4	14006	8	US-10-311-455-1931	Sequence 1931, Appl
17	63.2	3.4	1509	7	US-10-333-002-15	Sequence 15, Appl

C	18	-63	3.3	822	9	US-10-363-345A-11443	Sequence 31444, A
C	19	63	3.3	822	9	US-10-363-345A-11444	Sequence 31444, A
C	20	63	3.3	822	9	US-10-363-345A-38409	Sequence 38409, A
C	21	63	3.3	822	9	US-10-363-345A-38410	Sequence 38410, A
C	22	63	3.3	822	10	US-10-363-483A-31443	Sequence 31443, A
C	23	63	3.3	822	10	US-10-363-483A-31444	Sequence 31444, A
C	24	63	3.3	822	10	US-10-363-483A-38409	Sequence 38409, A
C	25	63	3.3	822	10	US-10-363-483A-38410	Sequence 38410, A
C	26	62.6	3.3	739	9	US-10-425-115-55500	Sequence 55020, A
C	27	62.4	3.3	1214	8	US-10-424-599-102083	Sequence 102083, A
C	28	61.8	3.3	9539	6	US-10-239-676-52	Sequence 52, App1
C	29	61.8	3.3	9539	7	US-10-240-445-54	Sequence 54, App1
C	30	61.6	3.3	516	3	US-09-960-352-5785	Sequence 5785, App
C	31	61.2	3.3	647	9	US-10-425-115-40312	Sequence 40312, A
C	32	61	3.3	1130	9	US-10-425-115-123399	Sequence 133399, A
C	33	60.6	3.2	1065	16	US-11-189-817-1	Sequence 1, App1
C	34	60.4	3.2	1039	9	US-10-425-115-56802	Sequence 56802, A
C	35	60.4	3.2	1243	9	US-10-425-115-172717	Sequence 172717, A
C	36	60	3.2	1121	9	US-10-425-115-115706	Sequence 115706, A
C	37	60	3.2	1297	9	US-10-425-115-67687	Sequence 67687, A
C	38	60	3.2	15674	7	US-10-311-455-336	Sequence 336, App
C	39	60	3.2	15674	7	US-10-240-448-30	Sequence 30, App1
C	40	59.8	3.2	5750	7	US-10-311-455-1981	Sequence 1981, App
C	41	59.4	3.2	5750	8	US-10-221-711A-433	Sequence 433, App
C	42	59.4	3.2	556	9	US-10-357-930-40082	Sequence 40082, A
C	43	59.4	3.2	556	9	US-10-357-930-40182	Sequence 40182, A
C	44	59.4	3.2	556	9	US-10-357-930-42124	Sequence 42124, A
C	45	59.4	3.2	556	9	US-10-357-930-43620	Sequence 43620, A

ALIGNMENTS

```

RESULT 1
US-10-661-809-12
: Sequence 12, Application US/10661809
: Publication No. US20040101919A1
: GENERAL INFORMATION:
: APPLICANT: HOOK, Magnus
: TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FR
: FILE REFERENCE: P07741US01/BAS
: CURRENT APPLICATION NUMBER: US/10/661,809
: PRIOR FILING DATE: 2003-09-15
: PRIOR APPLICATION NUMBER: 60/410303
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 13
: LENGTH: 1881
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
US-10-661-809-12

```

Query Match	100.0%;	Score 1881;	DB 8;	Length 1881;
-------------	---------	-------------	-------	--------------

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0

[illegible]

241 AACGTGAGAAAGAAATTTTACGAGCAACGAGGCGCAAGCGTGAATGACGTA 300
241 AACGTGAGAAAGAAATTTTACGAGCAACGAGGCGCAAGCGTGAATGACGTA 300
301 CAAGCTGTCGAAAGTTTAACTCTGGGAAAACCTGTGCTCAAGAAACCAACGATGCA 360
301 CAAGCTGTCGAAAGTTTAACTCTGGGAAAACCTGTGCTCAAGAAACCAACGATGCA 360
361 GGGAAATGTCAGTTCCTTAACCTTAAACCAAAATGTTAAAGATGCAATGATAC 420
361 GGGAAATGTCAGTTCCTTAACCTTAAACCAAAATGTTAAAGATGCAATGATAC 420
421 AAAAGAAACCAAAAGAGGTGATGCTGCTAGCAATATGCTGCTGCTGCTGCTGCT 480
421 AAAAGAAACCAAAAGAGGTGATGCTGCTAGCAATATGCTGCTGCTGCTGCTGCT 480
481 TACGAATGATCAAGCAACAGATGCTCTATTAATATGAAACAGAAATTAAGCGGT 540
481 TACGAATGATCAAGCAACAGATGCTCTATTAATATGAAACAGAAATTAAGCGGT 540
541 GTTCATATTTATCTTAAAAATGCTGAGCAATGATGATGTTTACATGTAAGAAAGTA 600
541 GTTCATATTTATCTTAAAAATGCTGAGCAATGATGATGTTTACATGTAAGAAAGTA 600
601 GGAACCTGCGAAAATGAAAGATTTAAATGCGGCAAGATTTGTTATTTTAAAGCGAAG 660
601 GGAACCTGCGAAAATGAAAGATTTAAATGCGGCAAGATTTGTTATTTTAAAGCGAAG 660
661 TCACGAGGACAGTAATATATATCCAGAGATCAAGATGATGATTAATATATATGCAAG 720
661 TCACGAGGACAGTAATATATATCCAGAGATCAAGATGATGATTAATATATATGCAAG 720
721 GATTAAGAAACAAAGCAAAACGCTTTTATTAATGCGGAAAAGTTAAATGAAATGCA 780
721 GATTAAGAAACAAAGCAAAACGCTTTTATTAATGCGGAAAAGTTAAATGAAATGCA 780
781 TTCACGAGGACAGTAATGAAACGAGGAAATTAACGTTAAATCTTAAAGGTGCTG 840
781 TTCACGAGGACAGTAATGAAACGAGGAAATTAACGTTAAATCTTAAAGGTGCTG 840
841 TATATTTTGAAGAAGTAATGCTCAATTAATGCAATTAATGAAATCAACAAA 900
841 TATATTTTGAAGAAGTAATGCTCAATTAATGCAATTAATGAAATCAACAAA 900
901 ACACCAATTTACATTTGAGCAAAACATCAACACCTGTTGAAAAAAACAGTCAAAATGAT 960
901 ACACCAATTTACATTTGAGCAAAACATCAACACCTGTTGAAAAAAACAGTCAAAATGAT 960
961 ACCTCTAAAGTTGATTAACCAACCAACGCTTAAGATGTAAGATGCGCAATGCGCA 1020
961 ACCTCTAAAGTTGATTAACCAACCAACGCTTAAGATGTAAGATGCGCAATGCGCA 1020
1021 AAAATTAATATCAATTTCTGTAATATTTCCATTTGGGATGCGACAAAAGAGCGAC 1080
1021 AAAATTAATATCAATTTCTGTAATATTTCCATTTGGGATGCGACAAAAGAGCGAC 1080
1081 GCTAATTAATACGTCAATTTCAATTTAGTGAATTAACATGATGCACTTAACTTTGAT 1140
1081 GCTAATTAATACGTCAATTTCAATTTAGTGAATTAACATGATGCACTTAACTTTGAT 1140
1141 AACGTGACCTTGGAGAGATGCTTATGCTTATATGAGGAGATCAAGATGCTGCT 1200
1141 AACGTGACCTTGGAGAGATGCTTATGCTTATATGAGGAGATCAAGATGCTGCT 1200
1201 GAAATTTATCAAGTGAAGCAAGCAAAATGCTGCTGCTGCTGCTTATCCAGGAT 1260
1201 GAAATTTATCAAGTGAAGCAAGCAAAATGCTGCTGCTGCTGCTTATCCAGGAT 1260
1261 ATTCTACGCTAACACGAGGCGGACACTTAAATTTGTTACTTATGCAATTTAAATGAA 1320
1261 ATTCTACGCTAACACGAGGCGGACACTTAAATTTGTTACTTATGCAATTTAAATGAA 1320

1321 AAAGCAGATCTACGAAGCTTTTAAATATGAGCGAATGTTGATTAACGTCATACGAC 1380
1321 AAAGCAGATCTACGAAGCTTTTAAATATGAGCGAATGTTGATTAACGTCATACGAC 1380
1381 GACCAACACCAACCAACTGTTGAAGTTGTAACAGTGGGAAAACGTTTATTAAGTCAT 1440
1381 GACCAACACCAACCAACTGTTGAAGTTGTAACAGTGGGAAAACGTTTATTAAGTCAT 1440
1441 GCGCATGTCAGACGACACCAAGCTTGGCGGGAAGCTTCTTGTGTCGCGATCAAAAC 1500
1441 GCGCATGTCAGACGACACCAAGCTTGGCGGGAAGCTTCTTGTGTCGCGATCAAAAC 1500
1501 AGCGACAGCAAAATTTATTTGAAATTCGATGAAAACAGAAAGCAGCAACTGGGTGAA 1560
1501 AGCGACAGCAAAATTTATTTGAAATTCGATGAAAACAGAAAGCAGCAACTGGGTGAA 1560
1561 ACAAAGCTGAAGCAACTTCTTTTACAACAACGCTGATGATGATGATGATGATGATG 1620
1561 ACAAAGCTGAAGCAACTTCTTTTACAACAACGCTGATGATGATGATGATGATGATG 1620
1621 CTTAAATACGCTACTATTTTATTAAGAAATCTGATGATGATGATGATGATGATG 1680
1621 CTTAAATACGCTACTATTTTATTAAGAAATCTGATGATGATGATGATGATGATG 1680
1681 ACAATCGATGATTTGCTGCTCAATGATGATGATGATGATGATGATGATGATGATG 1740
1681 ACAATCGATGATTTGCTGCTCAATGATGATGATGATGATGATGATGATGATGATG 1740
1741 TCACGAGAAAGATGACCAACCAACCAAGGATCTTACCTTCAACAGGTGGCAAGGA 1800
1741 TCACGAGAAAGATGACCAACCAACCAAGGATCTTACCTTCAACAGGTGGCAAGGA 1800
1801 ATCTACGTTTCTTATGAGAGGCGGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1801 ATCTACGTTTCTTATGAGAGGCGGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1861 AGACGTAGAAAAAGAAATGCT 1881
1861 AGACGTAGAAAAAGAAATGCT 1881

RESULT 2
US-09-071-035-217
Sequence 217, Application US/09071035
Publication No. US20020045737A1
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 1983 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-071-035-217

Query Match 100.0%; Score 1881; DB 3; Length 1983;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAAGCATTTAAAAAGTTGGTACACCGTTAGTACCTGTGTACTATTTTGGCAGTT 60
DB 100 ATGAAGCATTTAAAAAGTTGGTACACCGTTAGTACCTGTGTACTATTTTGGCAGTT 159
QY 61 TTCAAGGTATTTAGGCAACAACCTGATTTGAGAGAAAATGGGAGCGCAG 120
DB 160 TTCAAGGTATTTAGGCAACAACCTGATTTGAGAGAAAATGGGAGCGCAG 219
QY 121 CTGATGATTCACAAAAGAAAATGACGATTTACGATCCGCTATTCCAAATAGCGG 180
DB 220 CTGATGATTCACAAAAGAAAATGACGATTTACGATCCGCTATTCCAAATAGCGG 279
QY 181 AAGAAATGAGCGATTTGATTAATATCAAGACTGGCAGATGTGACGTTTGTATTAT 240
DB 280 AAGAAATGAGCGATTTGATTAATATCAAGACTGGCAGATGTGACGTTTGTATTAT 339
QY 241 AACGTGCAAGCAAGATTTTACGAGCAAGAGGCGGCAAGCGTTGATGACAGTAA 300
DB 340 AACGTGCAAGCAAGATTTTACGAGCAAGAGGCGGCAAGCGTTGATGACAGTAA 399
QY 301 CAAGCTGTCCAAAGTTTAACTCTGAGAAACCTGTGCTCAAGAAACAACCGATGCAAT 360
DB 400 CAAGCTGTCCAAAGTTTAACTCTGAGAAACCTGTGCTCAAGAAACAACCGATGCAAT 459
QY 361 GGGATGTCACTGTTCAGTTACTTAAAAAACAATGTAAAGATGCACTGTATACANT 420
DB 460 GGGATGTCACTGTTCAGTTACTTAAAAAACAATGTAAAGATGCACTGTATACANT 519
QY 421 AAGAAAGAACCAAAAGAGGTGTAGTGTGCTGCTAGCAATATGATGAGCGTTCACAGT 480
DB 520 AAGAAAGAACCAAAAGAGGTGTAGTGTGCTGCTAGCAATATGATGAGCGTTCACAGT 579
QY 541 TACGAATGATCAACCAACAGATGTTCTTAATAATGAAACAGAAATTAAGCGTT 540
DB 580 TACGAATGATCAACCAACAGATGTTCTTAATAATGAAACAGAAATTAAGCGTT 639
QY 541 GTTCATATTATCTTAATAATGTGTGACCAATGATGTTTCACTGAAAAAGTA 600
DB 640 GTTCATATTATCTTAATAATGTGTGACCAATGATGTTTCACTGAAAAAGTA 699
QY 601 GGAATGCTGAAAATGAGAGATTAATGAGCGCAGAAATTTGTAATTTCTTAAAGCAAGG 660
DB 700 GGAATGCTGAAAATGAGAGATTAATGAGCGCAGAAATTTGTAATTTCTTAAAGCAAGG 759
QY 661 TCACCAAGCACAAGTAATATATCAAGAGATCAAGATGATTAATATCATGACAAAG 720
DB 760 TCACCAAGCACAAGTAATATATCAAGAGATCAAGATGATTAATATCATGACAAAG 819
QY 721 GATAAAGAACAGCAAAACGCTTTATATCTGAGAAAAGTTATGAAATTTGGCAAAATGAT 780
DB 820 GATAAAGAACAGCAAAACGCTTTATATCTGAGAAAAGTTATGAAATTTGGCAAAATGAT 879
QY 781 TTCAAGAGAGCAGAAATGAAACGGAGAAATTAACAGTTAAATCTTGAAGTTGCTCG 840
DB 880 TTCAAGAGAGCAGAAATGAAACGGAGAAATTAACAGTTAAATCTTGAAGTTGCTCG 939
QY 841 TATATTTTGAAGAGATTAAGCTCCAAATATGAGAAATTAATGAAATCAACAAA 900
DB 940 TATATTTTGAAGAGATTAAGCTCCAAATATGAGAAATTAATGAAATCAACAAA 999

```

```

QY 901 AACACATTTACATTTGAAGCAAAACATCAACACCTGTTGAAAAAACAAGTCAAAATGAT 960
DB 1000 AACACATTTACATTTGAAGCAAAACATCAACACCTGTTGAAAAAACAAGTCAAAATGAT 1059
QY 961 ACCCTTAAAGTTGATTAACCAACACCAAGCTTGAATGTAAGATGTCGAATTTGGCGAA 1020
DB 1060 ACCCTTAAAGTTGATTAACCAACACCAAGCTTGAATGTAAGATGTCGAATTTGGCGAA 1119
QY 1021 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGATTCAGACAAAGAGCGAC 1080
DB 1120 AAAATTTAATATCAAAATTTCTGTAAATATTCATTTGGGATTCAGACAAAGAGCGAC 1179
QY 1081 GCTAATTAATATGTAATTTGATGTTAAACATGATGAGCGCTTAATCTTTGAT 1140
DB 1180 GCTAATTAATATGTAATTTGATGTTAAACATGATGAGCGCTTAATCTTTGAT 1239
QY 1141 AACGTGACTTCTGAGAGATATGCTATGCTATATGATGAGGATATCAGTATTCCTCT 1200
DB 1240 AACGTGACTTCTGAGAGATATGCTATGCTATATGATGAGGATATCAGTATTCCTCT 1299
QY 1201 GAAAATTTATCAAGTATGAAACAGCAAAATGCTTCACTGTGCGGTTAATCAAGCTAT 1260
DB 1300 GAAAATTTATCAAGTATGAAACAGCAAAATGCTTCACTGTGCGGTTAATCAAGCTAT 1359
QY 1261 ATTCTAGGCTTAACCAAGCGGCAACCTAAATTCGTTACTTTATGATTAATGAA 1320
DB 1360 ATTCTAGGCTTAACCAAGCGGCAACCTAAATTCGTTACTTTATGATTAATGAA 1419
QY 1321 AAGACGATCTTACGAAAGGCTTTTAAAAATGAGCGAAATGTTGATATCGGTATACCGAC 1380
DB 1420 AAGACGATCTTACGAAAGGCTTTTAAAAATGAGCGAAATGTTGATATCGGTATACCGAC 1479
QY 1381 GACCAACACCAACCACTGTGAAGTTGTGACAGGTGGAAAAGTTTCAATTAAGTCAT 1440
DB 1480 GACCAACACCAACCACTGTGAAGTTGTGACAGGTGGAAAAGTTTCAATTAAGTCAT 1539
QY 1441 GCGGATGTGACAGCAACAGGCTTGGCGGAGGCTTCTTGTGTCGCGATCAAAAC 1500
DB 1540 GCGGATGTGACAGCAACAGGCTTGGCGGAGGCTTCTTGTGTCGCGATCAAAAC 1599
QY 1501 AGCGACACAGCAAAATTTTGAATAATGATGAAACAGAAAGACAGCACTTGGGTGAA 1560
DB 1600 AGCGACACAGCAAAATTTTGAATAATGATGAAACAGAAAGACAGCACTTGGGTGAA 1659
QY 1561 ACAAAAGCTGAAGCACTACTTTTACAACAACGCTGATGATGATGATATCACAGG 1620
DB 1660 ACAAAAGCTGAAGCACTACTTTTACAACAACGCTGATGATGATGATATCACAGG 1719
QY 1621 CTTAAATACGTAACCTATTAATTAAGAAAGAACTGAGCTCCGATGATTAATGCTTTGTA 1680
DB 1720 CTTAAATACGTAACCTATTAATTAAGAAAGAACTGAGCTCCGATGATTAATGCTTTGTA 1779
QY 1681 ACAATTCGATTTGAATTTGTGTCAATGAAACATGATGGAACAACAGAAACCTAGTT 1740
DB 1780 ACAATTCGATTTGAATTTGTGTCAATGAAACATGATGGAACAACAGAAACCTAGTT 1839
QY 1741 TCACCAAGAAAAGTACCAACAAACAAGGTAACCTTCAACAGGTGGCAAGGA 1800
DB 1840 TCACCAAGAAAAGTACCAACAAACAAGGTAACCTTCAACAGGTGGCAAGGA 1899
QY 1801 ATCTACGTTTCTTAAGAAAGGCGCAGTCTTGTCACTTAATTTGAGAGGTCACTTGTCT 1860
DB 1900 ATCTACGTTTCTTAAGAAAGGCGCAGTCTTGTCACTTAATTTGAGAGGTCACTTGTCT 1959
QY 1861 AGACGTAGAAAAGAAATGCT 1881
DB 1960 AGACGTAGAAAAGAAATGCT 1980

```

RESULT 3
US-10-206-576-217
; Sequence 217, Application US/10206576

Publication No. US20030017495A1
GENERAL INFORMATION:
APPLICANT: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/206,576
FILING DATE: 29-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369PDI1
INFORMATION FOR SEQ ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 1983 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 217:
US-10-206-576-217
Query Match 100.0%; Score 1881; DB 6; Length 1983;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGCAATTAATAAAGTTTGGTACACCGTTAGTACCTTGTACTAATTTTGCACCTT 60
DB 100 ATGAAGCAATTAATAAAGTTTGGTACACCGTTAGTACCTTGTACTAATTTTGCACCTT 159
QY 61 TTCAACAAGTGTATTTAGGACAAACAACCTGCAATTTGCAGAAATAAGGAGACGACAG 120
DB 160 TTCAACAAGTGTATTTAGGACAAACAACCTGCAATTTGCAGAAATAAGGAGAGCGACAG 219
QY 121 CTCGGATTTCAATAAAGAAATGAACGATTTACCGATCCGCTTATCAAAATAGCGGAG 180
DB 220 CTCGGATTTCAATAAAGAAATGAACGATTTACCGATCCGCTTATCAAAATAGCGGAG 279
QY 181 AAAGAAATGAGCGAGTTGATTAATATCAAGACCTGCGAGATGTGACGTTTATGATTTAT 240
DB 280 AAAGAAATGAGCGAGTTGATTAATATCAAGACCTGCGAGATGTGACGTTTATGATTTAT 339
QY 241 AACGTGACGAAGAAATTTTACGACGACGAGCGGCGCAAGCGCTTATGACAGTTAA 300
DB 340 AACGTGACGAAGAAATTTTACGACGACGAGCGGCGCAAGCGCTTATGACAGTTAA 399
QY 301 CAAGCGTCAAAAGTTTAACTCTCGGGAACCGTTGCTCAAGGAACGAGATGCAAT 360
DB 400 CAAGCGTCAAAAGTTTAACTCTCGGGAACCGTTGCTCAAGGAACGAGATGCAAT 459
QY 361 GGGAAATGTCATGTTCAAGTTACCTAATAAATAAATGTAAGATGCAATGATACCAAT 420

DB 460 GGGAAATGTCATGTTCAAGTTACCTAATAAATAAATGTAAGATGCAATGATACCAAT 519
QY 421 AAAGAAATGAGCGAGTGTAGTTCTGCTAGCAATTAAGTGTGCGCTTCCAGTT 480
DB 520 AAAGAAATGAGCGAGTGTAGTTCTGCTAGCAATTAAGTGTGCGCTTCCAGTT 579
QY 481 TACGAATGATCAAGCAACAGATGTTCTTAATAATATGAAACGAAGAAATTAACGTT 540
DB 580 TACGAATGATCAAGCAACAGATGTTCTTAATAATATGAAACGAAGAAATTAACGTT 639
QY 541 GTTCATATTTATCTTAATAATGTGTAGCCAAATGATGTTTACATGTAAAAAAGTA 600
DB 640 GTTCATATTTATCTTAATAATGTGTAGCCAAATGATGTTTACATGTAAAAAAGTA 699
QY 601 GGAATCTGTGAAAAATGAAGATTAATGCGGAGAAATTTGATTTCTAAAGCGAAGC 660
DB 700 GGAATCTGTGAAAAATGAAGATTAATGCGGAGAAATTTGATTTCTAAAGCGAAGC 759
QY 661 TCACGAGGACAGTAAATATATCCAGAGATCAAGAATGATTTATATACATGACCAAG 720
DB 760 TCACGAGGACAGTAAATATATCCAGAGATCAAGAATGATTTATATACATGACCAAG 819
QY 721 GATTAAGAAACAGCAAAACGCTTTATTACTGGGAAAAGTTATGAAATGGCGAAAATGAT 780
DB 820 GATTAAGAAACAGCAAAACGCTTTATTACTGGGAAAAGTTATGAAATGGCGAAAATGAT 879
QY 781 TTCAAGAAAGCAAGAAATGGAACCGGGAATTAACGTTAAATACTGAGTGTGCTCG 840
DB 880 TTCAAGAAAGCAAGAAATGGAACCGGGAATTAACGTTAAATACTGAGTGTGCTCG 939
QY 841 TATATTTAGAAAGATTAAGTAAAGCTTCAATATATGCAATTAATTAATGAATCAACAAA 900
DB 940 TATATTTAGAAAGATTAAGTAAAGCTTCAATATATGCAATTAATTAATGAATCAACAAA 999
QY 901 ACACCAATTAACATTTGAACCAACATTAACACCTGTGAAAAACGTCAAAAATGAT 960
DB 1000 ACACCAATTAACATTTGAACCAACATTAACACCTGTGAAAAACGTCAAAAATGAT 1059
QY 961 ACCCTTAAGTTGAATAAACAACACAGCTTAGATGTTAAAGATGTCGCAATTCGCGAA 1020
DB 1060 ACCCTTAAGTTGAATAAACAACACAGCTTAGATGTTAAAGATGTCGCAATTCGCGAA 1119
QY 1021 AAAATTAATATCAATTTTCTGTAATATATTCATTTGGGGAATTCAGACAAAGAGCGAC 1080
DB 1120 AAAATTAATATCAATTTTCTGTAATATATTCATTTGGGGAATTCAGACAAAGAGCGAC 1179
QY 1081 GCTAATTAATATCGTCAATTTCAATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 1140
DB 1180 GCTAATTAATATCGTCAATTTCAATTTAGTTGATTAACATGATGACGCTTAACTTTGAT 1239
QY 1141 AACGTGACTTCGAGAGATATGCTTAATGCTTAATGATGAGGAGATACAGTATGCTCCT 1200
DB 1240 AACGTGACTTCGAGAGATATGCTTAATGCTTAATGATGAGGAGATACAGTATGCTCCT 1299
QY 1201 GAAATTTATCAAGTACGTAACAACAAATGAGCTTCACTGTCGCCGTTAATTCACGAT 1260
DB 1300 GAAATTTATCAAGTACGTAACAACAAATGAGCTTCACTGTCGCCGTTAATTCACGAT 1359
QY 1261 ATTCCTAGCTTAACACAGCGGCGCACTAAATTCGTTTACTTTATGATTAATGAA 1320
DB 1360 ATTCCTAGCTTAACACAGCGGCGCACTAAATTCGTTTACTTTATGATTAATGAA 1419
QY 1321 AAAGAGATTCCTAAGAAAGGCTTTAAATAAGGGAATGTTGATTAAGGTCATCCGAC 1380
DB 1420 AAAGAGATTCCTAAGAAAGGCTTTAAATAAGGGAATGTTGATTAAGGTCATCCGAC 1479
QY 1381 GACCAAAACACACCAACTGTTGAAGTTGACAGGTGGGAAAACGTTTCAATTAAGCTGAT 1440
DB 1480 GACCAAAACACACCAACTGTTGAAGTTGACAGGTGGGAAAACGTTTCAATTAAGCTGAT 1539
QY 1441 GCGGATGTGACAGGACCAACAGCTTGGCGGGAAGCTTCTTTGTCGTCGATCAAAAC 1500
DB 1540 GCGGATGTGACAGGACCAACAGCTTGGCGGGAAGCTTCTTTGTCGTCGATCAAAAC 1599

Db 1060 ACCTTAAGTTGATTAACCAACCAAGCTTATGATGTAAGATGTGCAATTTGGCCAA 1119
QY 1021 AAAATTAATCAAAATTTCTGTAAATATTCATTTGGGGATTCAGACAAAGAGCCAC 1080
Db 1120 AAAATTAATCAAAATTTCTGTAAATATTCATTTGGGGATTCAGACAAAGAGCCAC 1179
QY 1081 GCTAATAAATAGTCAAAATTTAGTATGATTAACATGAGCCCTTAATCTTTGAT 1140
Db 1180 GCTAATAAATAGTCAAAATTTAGTATGATTAACATGAGCCCTTAATCTTTGAT 1239
QY 1141 AACGTACTCTTGAGAGATATCTTATGCTTATATGATGGGATATCAAGTATGCTCCT 1200
Db 1240 AACGTACTCTTGAGAGATATCTTATGCTTATATGATGGGATATCAAGTATGCTCCT 1299
QY 1201 GAAATTTATCAAGTATGATCAACAGCAATGCTTCACTGTGCGCTTATATCAGCGTAT 1260
Db 1300 GAAATTTATCAAGTATGATCAACAGCAATGCTTCACTGTGCGCTTATATCAGCGTAT 1359
QY 1261 ATTCTAGCGTAAACCGAGGGGGACACTAAATTCGTTACTTTATGATTAATGAA 1320
Db 1360 ATTCTAGCGTAAACCGAGGGGGACACTAAATTCGTTACTTTATGATTAATGAA 1419
QY 1321 AAAGCAGATCTTACGAAAGGCTTTAAATATGAGCGAATGTTATACGTCATACCGAC 1380
Db 1420 AAAGCAGATCTTACGAAAGGCTTTAAATATGAGCGAATGTTATACGTCATACCGAC 1479
QY 1381 GACCAACACCAACCACTGTGATGATGACAGGTGGGAAACGTTTCAATTAAGTCAT 1440
Db 1480 GACCAACACCAACCACTGTGATGATGACAGGTGGGAAACGTTTCAATTAAGTCAT 1539
QY 1441 GGGGATGTGACAGCGACCAAGCTTTGGGGGAGCTTCTTTGCTGCTGATCAAAAC 1500
Db 1540 GGGGATGTGACAGCGACCAAGCTTTGGGGGAGCTTCTTTGCTGCTGATCAAAAC 1599
QY 1501 AGCGACACGCAAAATTTATGAAATCGATGAAACACGAAAGCAAGCACTTGGGTGAA 1560
Db 1600 AGCGACACGCAAAATTTATGAAATCGATGAAACACGAAAGCAAGCACTTGGGTGAA 1659
QY 1561 ACAAAAGCTGAAGCAACTCTTTTACAAACAAGCGCTGATGATTAATGATATCACAGG 1620
Db 1660 ACAAAAGCTGAAGCAACTCTTTTACAAACAAGCGCTGATGATTAATGATATCACAGG 1719
QY 1621 CTTAATAAGTATCTAATTTATTAAGAAAGCTGATGCTCGATGATTAATGCTTTGTA 1680
Db 1720 CTTAATAAGTATCTAATTTATTAAGAAAGCTGATGCTCGATGATTAATGCTTTGTA 1779
QY 1681 ACAAAATCGGATTTGAATTTGTGTCAATGAAATCATATGACCAACAGAAAACTTATGTT 1740
Db 1780 ACAAAATCGGATTTGAATTTGTGTCAATGAAATCATATGACCAACAGAAAACTTATGTT 1839
QY 1741 TCACCAAGAAAGTATCAACCAACCAAGGTAACCTTTCATTCACAGGTGGCAAGAA 1800
Db 1840 TCACCAAGAAAGTATCAACCAACCAAGGTAACCTTTCATTCACAGGTGGCAAGAA 1899
QY 1801 ATCTAGTTTACTTATGAGAGTGGCGAGCTTGTCTACTATGACAGAGTCTACTTTGCT 1860
Db 1900 ATCTAGTTTACTTATGAGAGTGGCGAGCTTGTCTACTATGACAGAGTCTACTTTGCT 1959
QY 1861 AGACGTAGAAAAAGAAATGCT 1881
Db 1960 AGACGTAGAAAAAGAAATGCT 1980

RESULT 5
US-09-070-927A-45
Sequence 45, Application US/09070927A
Patent No. US20020120116A1

GENERAL INFORMATION:

APPLICANT: Charles A. Kunsch
Patricia J. Dillon

Steven Barash

TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 15614 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-070-927A-45
Query Match 100.0%; Score 1881; DB 3; Length 15614;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACTTGTCTAATTTTGGCACTT 60
Db 12164 ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACTTGTCTAATTTTGGCACTT 12223
QY 61 TTCACAACTGTATTAGGACAAACAATGCAATTTGCAAGAAATGGGAGAGCGCACAG 120
Db 12224 TTCACAACTGTATTAGGACAAACAATGCAATTTGCAAGAAATGGGAGAGCGCACAG 12283
QY 121 CTCGTGATTCACAAAAAAGAAATGACGGAATTTACAGATCCGCTTAATCAAAATGCGGG 180
Db 12284 CTCGTGATTCACAAAAAAGAAATGACGGAATTTACAGATCCGCTTAATCAAAATGCGGG 12343
QY 181 AAAAGAAATGAGGAGTTGATTAATATATCAAGGACTGGCAGATGTGACGTTTATAT 240
Db 12344 AAAAGAAATGAGGAGTTGATTAATATATCAAGGACTGGCAGATGTGACGTTTATAT 12403
QY 241 AACGTAGCAAGAAATTTTACGAGCAACGAGCGGAGCGCAAGCGTTGATGACCTTAA 300
Db 12404 AACGTAGCAAGAAATTTTACGAGCAACGAGCGGAGCGCAAGCGTTGATGACCTTAA 12463
QY 301 CAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGTGCTCAAGGAAACCAAGATGCAAT 360
Db 12464 CAAGCTGTCCAAAGTTTAACTCTGGGAAACCTGTGCTCAAGGAAACCAAGATGCAAT 12523
QY 361 GGGATGTCACTGTCACTTACCTTAAATAAATAAATGTTAAAGATGCAAGTATACCAT 420
Db 12524 GGGATGTCACTGTCACTTACCTTAAATAAATAAATGTTAAAGATGCAAGTATACCAT 12583
QY 421 AAAGAAAGCAAAAAAGAGGTGTAGTCTGTCTAGCAATATGCTGTGCGTTCCAGATT 480

QY 96 AGAAGAAATGGGGAGAGCGCAGCTGATTCACAAAAAGAAATGACGATTTACC 155
Db 1 AGAAGAAATGGGGAGAGCGCAGCTGATTCACAAAAAGAAATGACGATTTACC 60
QY 156 AGATCCGCTTATTCAAAAATGCGGAGAAAGAAATGAGCGAGTTGATTAATCAAGACT 215
Db 61 AGATCCGCTTATTCAAAAATGCGGAGAAAGAAATGAGCGAGTTGATTAATCAAGACT 120
QY 216 GGCACATGAGCGTTAGTATTTATTAACGTGACGAACGAAATTTTCAGACCAAGCGGC 275
Db 121 GGCACATGAGCGTTAGTATTTATTAACGTGACGAACGAAATTTTCAGACCAAGCGGC 180
QY 276 AGGCCAAGCGTTGATGACGCTAAACAGCTGCCAAGTTTAACTCTCGGAAAAGCTGT 335
Db 181 AGGCCAAGCGTTGATGACGCTAAACAGCTGCCAAGTTTAACTCTCGGAAAAGCTGT 240
QY 336 TGCTCAAGAACCCACCGATGCAATGGGAATGTCACTGTTCACTTAACTTAAACAAA 395
Db 241 TGCTCAAGAACCCACCGATGCAATGGGAATGTCACTGTTCACTTAACTTAAACAAA 300
QY 396 TGGTAAAGATGCGATGATACCAATTAAGAAACCAAAAGAGGGTGTGCTGCTAC 455
Db 301 TGGTAAAGATGCGATGATACCAATTAAGAAACCAAAAGAGGGTGTGCTGCTAC 360
QY 456 GAATATGATGATGCGCTCCAGTTTACGAAATGATCAAGCAAGAGATGCTTCTATTA 515
Db 361 GAATATGATGATGCGCTCCAGTTTACGAAATGATCAAGCAAGAGATGCTTCTATTA 420
QY 516 ATATGGAACAGAAAGATTAGCGGTGTTCAATTTATCTTAAATATGCTGACCAATGA 575
Db 421 ATATGGAACAGAAAGATTAGCGGTGTTCAATTTATCTTAAATATGCTGACCAATGA 480
QY 576 TGGTATGTTTACATGGAAGAAAGTGAAGATGCTGAAATGGAAGTTTAAATGGGCGACA 635
Db 481 TGGTATGTTTACATGGAAGAAAGTGAAGATGCTGAAATGGAAGTTTAAATGGGCGACA 540
QY 636 ATTTGTTATTTCTTAAAGCGAAGGCTCAACAGCAAGTAAATATATCCAGAGATCA 695
Db 541 ATTTGTTATTTCTTAAAGCGAAGGCTCAACAGCAAGTAAATATATCCAGAGATCA 600
QY 696 AGATGATTTATATACATGACCAACGATTAAGAACAGCAAAACGCTTATTAATCTGGGAA 755
Db 601 AGATGATTTATATACATGACCAACGATTAAGAACAGCAAAACGCTTATTAATCTGGGAA 660
QY 756 AAGTATGAAATTTGGGAAAATGATTTTCACAGAGCAGAAATGGAACGGAGAAATTAAC 815
Db 661 AAGTATGAAATTTGGGAAAATGATTTTCACAGAGCAGAAATGGAACGGAGAAATTAAC 720
QY 816 AGTTAAAAATCTTGAGGTTGTTGCTATATTTTGAAGAGTAAAGCTCCAAATTAATGC 875
Db 721 AGTTAAAAATCTTGAGGTTGTTGCTATATTTTGAAGAGTAAAGCTCCAAATTAATGC 780
QY 876 AGAATTAATTTGAATCAAAACAAAACACCATTTTACATTTGAAGCAACATCAAAACCC 935
Db 781 AGAATTAATTTGAATCAAAACAAAACACCATTTTACATTTGAAGCAACATCAAAACCC 840
QY 936 TGTGAAAAAAGCAGTCAAAAATGATACCTCTTAAAGTTGATTAACCAACCAAGCTTGA 995
Db 841 TGTGAAAAAAGCAGTCAAAAATGATACCTCTTAAAGTTGATTAACCAACCAAGCTTGA 900
QY 996 TGGTAAAGATGCGCAATTTGGGAAAATTAATTAATCAAAATTTCTGTAATTAATTCATT 1055
Db 901 TGGTAAAGATGCGCAATTTGGGAAAATTAATTAATCAAAATTTCTGTAATTAATTCATT 960
QY 1056 GGGGATTTGACACAAAGAGGCGGCTAATAATTAATTCGCAATTTCAATTTAGTTGATA 1115
Db 961 GGGGATTTGACACAAAGAGGCGGCTAATAATTAATTCGCAATTTCAATTTAGTTGATA 1020
QY 1116 ACATGATGACGCTTAACTTTTGAATTAACGTCTTGAGAGATAGCTTATGCTTATTA 1175
Db 1021 ACATGATGACGCTTAACTTTTGAATTAACGTCTTGAGAGATAGCTTATGCTTATTA 1080
QY 1176 TGATGGGATACAGTGAATTTCTCTGAAAAATTAATCAAGTGAACAAAGCAATGGCTT 1235

Db 1081 TGATGGGATACAGGATTTGCTCTGAAAATTAATCAAGTGAACAAAGCAATGGCTT 1140
QY 1236 CACTGTCCCGCTTAAATCCAGGATATATCTTAACGTAAACACAGGCGGCACTAAATTT 1295
Db 1141 CACTGTCCCGCTTAAATCCAGGATATATCTTAACGTAAACACAGGCGGCACTAAATTT 1200
QY 1296 CGTTTACTTTATGCAATTTAAATGAANAAGCAGATCTTCAAGAAAGGCTTTAAATATGAGGC 1355
Db 1201 CGTTTACTTTATGCAATTTAAATGAANAAGCAGATCTTCAAGAAAGGCTTTAAATATGAGGC 1260
QY 1356 GAATGTTGATAACGGTCAATACCGACGACCAACACACCAACTGTTGAAGTTTGACAGG 1415
Db 1261 GAATGTTGATAACGGTCAATACCGACGACCAACACACCAACTGTTGAAGTTTGACAGG 1320
QY 1416 TGGGAAGGTTTCAATTAAGTCGATGAGGAGATGACAGCACAAGAGCTTGGCGGGAGC 1475
Db 1321 TGGGAAGGTTTCAATTAAGTCGATGAGGAGATGACAGCACAAGAGCTTGGCGGGAGC 1380
QY 1476 TTCCTTTGCTCGCTGATCAAAAACAGCGACACAGCAATTAATTTGAANAATGATGAAC 1535
Db 1381 TTCCTTTGCTCGCTGATCAAAAACAGCGACACAGCAATTAATTTGAANAATGATGAAC 1440
QY 1536 AACGAAGCAGCACTTGGGTGAAAACAAAGCTGAAGCACTTTTAAACAAACGCGC 1595
Db 1441 AACGAAGCAGCACTTGGGTGAAAACAAAGCTGAAGCACTTTTAAACAAACGCGC 1500
QY 1596 TGAATGATTAAGTTGATATCAAGGGCTTAAATAGGCTTATTAATTAAGAAAGATGCT 1655
Db 1501 TGAATGATTAAGTTGATATCAAGGGCTTAAATAGGCTTATTAATTAAGAAAGATGCT 1560
QY 1656 AGCTCTGATGATTAATGCTTTGTTAAACAATCGAATGAATTTGTGCTCAATGAACAATC 1715
Db 1561 AGCTCTGATGATTAATGCTTTGTTAAACAATCGAATGAATTTGTGCTCAATGAACAATC 1620
QY 1716 ATATGCGCAACAGAAAACCTTAGTTTCAACGAAAAGTACCAACAAACAGAGTAC 1775
Db 1621 ATATGCGCAACAGAAAACCTTAGTTTCAACGAAAAGTACCAACAAACAGAGTAC 1680
QY 1776 CTTACT 1782
Db 1681 CTTACT 1687

RESULT 7
US-10-206-576-219
; Sequence 219, Application US/10206576
; Publication No. US20030017495A1
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16

APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D1
SEQUENCE CHARACTERISTICS:
LENGTH: 1687 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 219:
US-10-206-576-219

Query Match 89.7%; Score 1687; DB 6; Length 1687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 AGAAGAAAATGGGGAGAGCGCACAGCTCGTGAATTCACAAAAGAAAATGACGATTTACC 155
DB 1 AGAAGAAAATGGGGAGAGCGCACAGCTCGTGAATTCACAAAAGAAAATGACGATTTACC 60
QY 156 AGATCCGCTTATTCAAAATAGCGGGAAAGAAATAGCGAGTTGATTAATATCAAGGACT 215
DB 61 AGATCCGCTTATTCAAAATAGCGGGAAAGAAATAGCGAGTTGATTAATATCAAGGACT 120
QY 216 GGCAGATGTCAGCTTAAATTAATTAACGTGACGAAAGAAATTTTACGACCAAGAGCGGC 275
DB 121 GGCAGATGTCAGCTTAAATTAATTAACGTGACGAAAGAAATTTTACGACCAAGAGCGGC 180
QY 276 AGGCGCAAGCGTTGATGACGCTAAACAGCTGCTCAAAAGTTTAACTCTGGGAAACTGT 335
DB 181 AGGCGCAAGCGTTGATGACGCTAAACAGCTGCTCAAAAGTTTAACTCTGGGAAACTGT 240
QY 336 TGCTCAAGGAACCAACGATGCAAAATGGGAAATGTCATGTTCACTTAACCTAAACAAACAA 395
DB 241 TGCTCAAGGAACCAACGATGCAAAATGGGAAATGTCATGTTCACTTAACCTAAACAAACAA 300
QY 396 TGGTAAAGATGTCAGTATACCATTAAGAAAGAAACCAAAAGAGGGTGTAGTGTCTAC 455
DB 301 TGGTAAAGATGTCAGTATACCATTAAGAAAGAAACCAAAAGAGGGTGTAGTGTCTAC 360
QY 456 GAATATGCTGTGCGGCTTCCAGTTTACGAAATGATCAAGCAACAGATGTTCTATTA 515
DB 361 GAATATGCTGTGCGGCTTCCAGTTTACGAAATGATCAAGCAACAGATGTTCTATTA 420
QY 516 AATATGGAACAGAAAGAAATAGCGGTTGTCATATTTATCCTAAACAAATGTGTGCAATTA 575
DB 421 AATATGGAACAGAAAGAAATAGCGGTTGTCATATTTATCCTAAACAAATGTGTGCAATTA 480
QY 576 TGGTATTTTACATGTAAGAAAGTGAAGACTGCTGAATAATGAAGATTAAATGGCGCAGA 635
DB 481 TGGTATTTTACATGTAAGAAAGTGAAGACTGCTGAATAATGAAGATTAAATGGCGCAGA 540
QY 636 ATTGTGTTATTTCTAAAAGGAGGCTCACAGGCAACAGTAAATATATCAAGAGTCA 695
DB 541 ATTGTGTTATTTCTAAAAGGAGGCTCACAGGCAACAGTAAATATATCAAGAGTCA 600
QY 696 AGATGATTTATTAATGACGAAACGTAAGAAAGAAACGTTTAAATTTACCTGGGAA 755
DB 601 AGATGATTTATTAATGACGAAACGTAAGAAAGAAACGTTTAAATTTACCTGGGAA 660
QY 756 AAGTTATGAAATTTGGCGAAATGATTTTACAGAGGAGAAATGGAACGGAGAAATTAAC 815
DB 661 AAGTTATGAAATTTGGCGAAATGATTTTACAGAGGAGAAATGGAACGGAGAAATTAAC 720
QY 816 AGTTAAATATCTTGAAGTTGTTGCTGATATTTTAAAGAAAGTAAAGCTTCAATTAATGC 875
DB 721 AGTTAAATATCTTGAAGTTGTTGCTGATATTTTAAAGAAAGTAAAGCTTCAATTAATGC 780

QY 876 AGAATTAATGAAATCAACAAACAAACACCATTAACATTTGAAGCAACATCAACACC 935
DB 781 AGAATTAATGAAATCAACAAACAAACACCATTAACATTTGAAGCAACATCAACACC 840
QY 936 TGTGAAAAACAGTCAAAAATGATACCTCTAAAGTTGATTAACCAACCAAGCTTAA 995
DB 841 TGTGAAAAACAGTCAAAAATGATACCTCTAAAGTTGATTAACCAACCAAGCTTAA 900
QY 996 TGTGAAAGATGAGCAATTTGGGAAAAATTAATATCAAAATTTCTGTAATATTCATT 1055
DB 901 TGTGAAAGATGAGCAATTTGGGAAAAATTAATATCAAAATTTCTGTAATATTCATT 960
QY 1056 GGGGATTCAGACAAAGAAAGGCGAGCTAATTAATACGTAATCAATTTAGTTGATTA 1115
DB 961 GGGGATTCAGACAAAGAAAGGCGAGCTAATTAATACGTAATCAATTTAGTTGATTA 1020
QY 1116 ACATGATGACGCTTAACCTTTGATTAACGTAACCTTCTGAGAGTATGCTTAATGCTTAA 1175
DB 1021 ACATGATGACGCTTAACCTTTGATTAACGTAACCTTCTGAGAGTATGCTTAATGCTTAA 1080
QY 1176 TGATGGGATACAGATTTGCTCTGAAATTAATCAAGTGAAGCAAGCAATGGCTT 1235
DB 1081 TGATGGGATACAGATTTGCTCTGAAATTAATCAAGTGAAGCAAGCAATGGCTT 1140
QY 1236 CACTGTCCGCTTAATCCAGCGTATATTCCTACGCTAACACAGGCGGACACTAAAT 1295
DB 1141 CACTGTCCGCTTAATCCAGCGTATATTCCTACGCTAACACAGGCGGACACTAAAT 1200
QY 1296 CGTTTACTTTATGCAATTAATGAAGAAAGAGATCCCTACGAAAGCTTTAAATGAGG 1355
DB 1201 CGTTTACTTTATGCAATTAATGAAGAAAGAGATCCCTACGAAAGCTTTAAATGAGG 1260
QY 1356 GAATGTTGAATACGGTCATACCGAGCAACCAACCAACGTTGAAGTTGTGACAG 1415
DB 1261 GAATGTTGAATACGGTCATACCGAGCAACCAACCAACGTTGAAGTTGTGACAG 1320
QY 1416 TGGGAAACGTTTCAATTAAGTGCATGCGATGACAGGACGACAAAGCTTGGCGGAGC 1475
DB 1321 TGGGAAACGTTTCAATTAAGTGCATGCGATGACAGGACGACAAAGCTTGGCGGAGC 1380
QY 1476 TTCCTTTGCTGCTGATCAAAACAGCGACACAGCAATTTATTTGAAATTCATGAAC 1535
DB 1381 TTCCTTTGCTGCTGATCAAAACAGCGACACAGCAATTTATTTGAAATTCATGAAC 1440
QY 1536 AACGAAAGCAACCTTGGGTGAAAAACAAACCTGAAGCACTTTTACCAACAGCGC 1595
DB 1441 AACGAAAGCAACCTTGGGTGAAAAACAAACCTGAAGCACTTTTACCAACAGCGC 1500
QY 1596 TGATGATTAATGATTAACAGGCTTAATACGTAACCTTAATTTTGAAGAACTGT 1655
DB 1501 TGATGATTAATGATTAACAGGCTTAATACGTAACCTTAATTTTGAAGAACTGT 1560
QY 1656 AGCTCTGATGATTAATGCTTTGTTAAACAAATGGAATGAATTTGTGTAATGAACAATC 1715
DB 1561 AGCTCTGATGATTAATGCTTTGTTAAACAAATGGAATGAATTTGTGTAATGAACAATC 1620
QY 1716 AATGCGACAAAGAAACCTTGTTCACCAAGAAAGTACCAACCAAGAGTAC 1775
DB 1621 AATGCGACAAAGAAACCTTGTTCACCAAGAAAGTACCAACCAAGAGTAC 1680
QY 1776 CTTACCT 1782
DB 1681 CTTACCT 1687

RESULT 8
US-10-912-362-219
Sequence 219, Application US/10912362
Publication No. US20050043528A1
GENERAL INFORMATION:
APPLICANT: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497

CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 14200 Shady Grove Road
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
OPERATING SYSTEM: Windows 2000
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/912,362
FILING DATE: 06-Aug-2004
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 10/206,576
FILING DATE: 2002-07-29
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369PID2
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 1687 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 219:
US-10-912-362-219
Query Match 89.7%; Score 1687; DB 9; Length 1687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

421 ATATGAAACAGAGAATTAGCGTTGTCATATTATCTTAATAAATGTGTAGCCAAATGA 480
576 TGGTATGTTTACATGTGAAAAAAGTATGAACTGCTGAAAAATGAAGATTAAATGGCGCAGA 635
481 TGGTAGTTTACATGTGAAAAAAGTATGAACTGCTGAAAAATGAAGATTAAATGGCGCAGA 540
636 ATTTGTTATTTTCTTAAGCGAAGGCTCACAGGCAAGTAAATATATCAAGAGTCA 695
541 ATTTGTTATTTTCTTAAGCGAAGGCTCACAGGCAAGTAAATATATCAAGAGTCA 600
636 AGATGATTTATATATCATGAGCAACCGAATTAAGAACAGCAAAACGCTTATTTACTGGGAA 755
601 AGATGATTTATATATCATGAGCAACCGAATTAAGAACAGCAAAACGCTTATTTACTGGGAA 660
756 AAGTTATGAATTTGGCGAAAAATGATTTTCAAGAACAGAGAAATGGAACGGGGAATTATAC 815
661 AAGTTATGAATTTGGCGAAAAATGATTTTCAAGAACAGAGAAATGGAACGGGGAATTATAC 720
816 AGTTAAAAATCTTGAGTTGGTTCGTATATTTTGAAGAAGTAAAGTCCAAATATATGC 875
721 AGTTAAAAATCTTGAGTTGGTTCGTATATTTTGAAGAAGTAAAGTCCAAATATATGC 780
876 AGAATTTAATTTGAAAAATCAACCAAAACCAATTTTCAATTGAAAGCAATCAAAACACC 935
781 AGAATTTAATTTGAAAAATCAACCAAAACCAATTTTCAATTGAAAGCAATCAAAACACC 840
936 TGTGAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 995
841 TGTGAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
996 TGTGAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1055
901 TGTGAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
1056 GGGGATTGACAGCAAAAGAGGCGCTAATTAATATGCTCAATTTAGTTGATTA 1115
961 GGGGATTGACAGCAAAAGAGGCGCTAATTAATATGCTCAATTTAGTTGATTA 1020
1116 ACATATGACAGCTTAACTTTTGAATAGGTGACTTCTGAGAGTATGCTTATGCTTATTA 1175
1021 ACATATGACAGCTTAACTTTTGAATAGGTGACTTCTGAGAGTATGCTTATGCTTATTA 1080
1176 TGATGGGATACAGGATTTGCTCTGAATAATTAATCAAGTGAACCAAGCAATGGGCT 1235
1081 TGATGGGATACAGGATTTGCTCTGAATAATTAATCAAGTGAACCAAGCAATGGGCT 1140
1236 CACTGTCGCGCTTAATTCAGCGTATATTCCTAAGCTTAAACCAAGCGGCACTAAATTT 1295
1141 CACTGTCGCGCTTAATTCAGCGTATATTCCTAAGCTTAAACCAAGCGGCACTAAATTT 1200
1296 CGTTTACTTTATGCAATTTTAAATGAATAAGCAGATCTTACGAAGGCTTTAAATATGAGGC 1355
1201 CGTTTACTTTATGCAATTTTAAATGAATAAGCAGATCTTACGAAGGCTTTAAATATGAGGC 1260
1356 GAATGTTGATACGGTCAATACCGAGACCAACCAACCACTGTTGAAGTTGTGACAGG 1415
1261 GAATGTTGATACGGTCAATACCGAGACCAACCAACCACTGTTGAAGTTGTGACAGG 1320
1416 TGGGAAAGGTTTCAATTAAGTGCATGGGATGTGACAGGCAACAAGCTTTGGCGGAGGC 1475
1321 TGGGAAAGGTTTCAATTAAGTGCATGGGATGTGACAGGCAACAAGCTTTGGCGGAGGC 1380
1476 TTCTTTTGTGCTGCGGTATCAAAACAGGCAACAGCAAAATTTTGAATAATGTATGAAC 1535
1381 TTCTTTTGTGCTGCGGTATCAAAACAGGCAACAGCAAAATTTTGAATAATGTATGAAC 1440
1536 AACGAAAGAGCACTTGGGATGAAAAAAGCTGAAGCACTACTTTTACAAACAGGC 1595
1441 AACGAAAGAGCACTTGGGATGAAAAAAGCTGAAGCACTACTTTTACAAACAGGC 1500
1596 TGATGATTTAGTTGATATCAAGGGCTTAAATACGATCTTATTTTGAAGAAACTGT 1655
1501 TGATGATTTAGTTGATATCAAGGGCTTAAATACGATCTTATTTTGAAGAAACTGT 1560

Qy	1556	AGCCTCCATGATATATGCTTGTTAA	CAAAATCGAGTTGAATTGGGTCATGAAACATC	1715
Db	1561	AGCTCTGATGATATATGCTTGTTAA	CAAAATCGAGTTGAATTGGGTCATGAAACATC	1620
Qy	1716	ATATGGCAACACAAAACCTTAGTTCC	CCGAAAAAGTACCAACAAACCAAAAGTAC	1775
Db	1621	ATATGGCAACACAAAACCTTAGTTCC	CCGAAAAAGTACCAACAAACCAAAAGTAC	1680
Qy	1776	CTTACTT	1782	
Db	1681	CTTACTT	1687	

```

RESULT 9
US-10-661-809-18.
: Sequence 18, Application US/10661809
: Publication No. US20040101919A1
: GENERAL INFORMATION:
: APPLICANT: HOOK, Magnus
: TITLE OF INVENTION: BIOINFORMATIC METHOD FOR IDENTIFYING SURFACE-ANCHORED PROTEINS FROM
: TITLE OF INVENTION: POSITIVE BACTERIA
: FILE REFERENCE: P07741US01/BAS
: CURRENT APPLICATION NUMBER: US/10/661,809
: CURRENT FILING DATE: 2003-09-15
: PRIOR APPLICATION NUMBER: 60/410303
: PRIOR FILING DATE: 2002-09-13
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 18
: LENGTH: 1878
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
: US-10-661-809-18

```

Query Match	53.9%	Score 1014.6	DB 8	Length 1878
Best Local Similarity	71.9%	Pred. No. 6.2e-194		
Matches 1342; Conservative	0	Mismatches 519	Indels 6	Gaps 1

QY	1	ATGAAAGCAATTA	AAAAAAGTTGGT	TACACCGCTAGTAC	TCTCTTGTAC	TAATTTGGC	CACT	60
Db	1	ATGAAAAATCAT	AAAAAAATTAAC	GTTATGT	TAGAGACTCTT	TCTTATTTTAC	CACTA	60
QY	61	TTCAAGATGAT	TTTAGGGA	CAACA	CTGCATTTTG	CAGAAAGAAA	TGGGAGAGCG	CAAG 120
Db	61	CTCAAAACAGCT	TCGGCGCA	AAAAAAGTTT	TCAGAGAGACAG	-----	CAGCTCAA	114
QY	121	CTGCGATTTCA	CAAAAAAGAAA	TGACGGA	TTTACACATCCGCT	TATTCAAAA	TAGCGGG	180
Db	115	GTCATCTTCA	TAAAAAGAAA	TACGTAGT	ATTCGCCATCTT	TATTC	CAAAACAGCGGG	174
QY	181	AAAGAAATGAC	GAGTTTGAT	TAAATAT	CAAGAC	CTGCGAGA	TGTGACGTTT	AGTATTTAT 240
Db	175	AAAGAAATGAC	GAAATTCGAT	CAATACCA	AGATTA	CCCAATTT	CAATTTTCAGTTAT	234
QY	241	AACGTGACGA	ACGAATTTT	TCGAGCA	CGACGCGCAG	CGCCAA	CGCTTGATG	CAAGCTTAA 300
Db	235	AACGTCACTC	AGAATTTTAT	TGCGCACA	CGAATTA	AGAGAGCTCG	CTGAGTCAG	CAACAAA 294
QY	301	CAAGCTGTCC	AAAGTTTAA	CTCTCGG	AGAAACCTGTGCT	CAAGAAC	CACCGATG	CAAT 360
Db	295	CAGACGATCC	AGCTTGTGAC	TCTCGGTAC	CCAGTTCTT	CAGGAA	GAGACAGATG	CGAT 354
QY	361	GGGAATGTCA	CTGTTCAGTT	ACTTAA	AAAAACAA	AATGTTAA	AGATGAGT	ATACCAT 420
Db	355	GGAAATGTCA	CTTATCTTTA	CTTAAAA	CAAAATGG	AAAGATG	AGCTAC	ACATC 414
QY	421	AAAGAAACCA	AAAGAGG	GTGTAGT	TCTGTACGA	ATATGTGTG	TGGCGTTCC	CAAGTT 480
Db	415	AAAGAAACCA	AAAGAGCGAG	GTGTACG	TCCGCGCA	AAATGTTT	TATGCTTTCC	CTGTA 474
QY	481	TACGAAATGAT	CAAGCAAGAT	GTGTTCC	TATTAATATG	AACAGAA	ATTAGCGTT	540

D	475	TATGAGATGATCAAA	CAAGCAATGCGCTTTA	TAATAAGGGG	CAGAAAGATGATAC	554	
Q	541	GTTCAATATTTATCT	TAATAATGTGTAGCCA	TGATGTAAGTTA	CAATGGAATAAAGTA	600	
D	535	ATCCATCTCTACCT	TAATAATCAAGTCGGTAA	TGATGGAACGT	TGAAGTTACAAAATC	594	
Q	601	GGAATCGCTGAATA	TGAAGGATTTAATG	CGCGAGAAATTTGTTA	TTTCTTAAGCGAAGC	660	
D	555	GGTATCGCGAAA	AGAAAGCATTAATG	AGCGAATTTATTTCT	TAAGAAAGAAAG	654	
Q	661	TCACAGGCA	CGTAATAATATAT	CCAAAGAGTC	CAAGATGATTAATAT	CAATGCAACG	720
D	655	ACACCAAGCTCA	AAAAAATATACCA	AAAGTGCAC	AGATGATTGTAT	CACTTGCAACAT	714
Q	721	GATTAAGAA	CAAGAAACGCTTTA	TTATCTGGGAAA	AGTTAAGAAATTTGG	CCAAAATGAT	780
D	715	GATCAACCA	AAAGCCAAACATT	TTCATTTACTGTGCAT	TTTATATGACATG	CGCAACATGAC	774
Q	781	TTCA	CAGAAGCAGAA	TGSAACGGGAGATTA	CAAGTTAAAAATCTT	GAGTTGGTTCG	840
D	775	TTTGC	CGAGGATCTAT	TGAAAAGGCCAGTTG	ATCGTTATATCATTTA	AGTTGGA	834
Q	841	TATATTTA	GAAAGATTAAGCT	CCAAATTAATG	CAATTAATGAAAAATCA	CAAA	900
D	835	TATATTTA	GAAAGATTAAGCT	CCGTATATGCG	GAATGATGAAAAAGCA	CAATC	894
Q	901	ACACATTTA	CAATTGAAGCAAT	CAACACCTGTTGA	AAAAACAGTCA	AAAAATGAT	960
D	895	ACGCTTTG	AGATCTCGGC	CAATATGCCAA	CAACGATGAAAGCA	CAATCAAAAATGAT	954
Q	961	ACCTTA	AGTTGATTAACA	CACCAAGCTT	GATGTTAAAGATG	GGCAATGGCGAA	1020
D	955	ACGTTA	AGTTGATTAACA	CACTCAATGGA	ATGAAAAAGATGTC	CAATCGGTGAA	1011
Q	1021	AAAATTA	ATATTCAAATTTCT	GTAAATATTCAT	TGGGATTTGC	CACAAAGAGCGAC	108
D	1015	AAAAATCA	ATATGAAATTTCT	GTCAATATCCAT	TATGATGCTGAT	TAAAGAAAGAACG	107
Q	1081	GCTATA	TAATATAGTCA	AAATTTCAATTTAG	TGATTAACAATGAT	GACCTTAACTTTTGAT	114
D	1075	CAAAAC	AGTACAAACATTTCA	AACTTATGATAT	CTCATGAGAGCTG	CTTTAATTTGAT	113
Q	1141	AACTG	ACTTTGAGAGATAT	GCTTATATGCGTTA	TATGATGGGATAT	CACTGATGCTCT	120
D	1135	AATGAT	CTTCAAGGAACGTAT	GCTTATATGAT	TGAAAAATTAAGAAATTC	GCACCA	119
Q	1201	GAAAAT	TATCAAGTACGTAA	CAAGCAATG	GCCTCACATGCGCGTTA	ATCCAGCTAT	126
D	1195	GTAATTTA	TTCTGTCACTG	AGCAACAGCGA	TTTCAAGGTTTCA	GTTTCAAGTTATCCGAAT	125
Q	1261	ATTCCT	CACTPAACACAG	CGGCACTATAAAT	TGTTACTTTATATG	CAATTTAAATGAA	132
D	1255	ATTCCT	CAATTAACCTCG	CGGATCAATGAA	ATTCGTTACTATATG	CAATTTGAACGA	131
Q	1321	AAACG	ATCTTAAGAAAGG	CTTTAAATAG	GGAATGTTGATTA	CGCTCATACCGAC	138
D	1315	AAACG	ATCTTAAGAAAGG	CTTTAAATAG	GGAATGTTGATTA	CGCTCATACCGAC	137
Q	1381	GACCA	AAACACCAACAT	CTGTGAAGTTG	ACAGGTGGAACGTTT	CAATTAAGTCAT	144
D	1375	GATCA	AAACACCAACG	CTGATCGATG	ATGCTTAACTGGG	CGAATGTTGTTAAATGAT	143
Q	1441	GGCA	TGTGACAGGCA	CAAGCCTTGG	CGGAGCTTCTTTG	TGTCCTGATCAAAAC	150
D	1435	GGTA	CGTTATCATCA	AGCAACCTTCTG	AGCAAAATTCGTCGTT	CGTATCAAGAT	149
Q	1501	AGCA	CAACAGCAAAAT	TATTTGAAATG	ATGAAACAAAGCAAG	CAACTTGGGTA	156
D	1495	AGTA	CAACAGCAAAAT	TATTTGATG	ATGAAACCAAAAG	CCGTCAGCTGGGATG	155
Q	1561	ACAA	AGCTGAACACTA	CTTTTACAC	CAACGCGCTAT	GATTAAGTTATCAAGG	162
D	1555	GCGAA	AGATCAGCAACG	TTTTTACAC	CAACAGTATCG	TTTTATGATGACAGT	161

```
OY 1621 CTTAAATACGTAACCTATTTAGAAAGAACTGTAAGTCTGATGATTAATGCTTTA 1680
DB 1615 CTAATAATATGCACTACTATCTGAAAGAAAGCAAGGCGCAAAAAATATGTCATTA 1674
OY 1681 ACAAAATCGATTTGAATTTGTGTCATGTAACATCATATGAGCAACAGAAAACTAGTT 1740
DB 1675 ACAAAACGATGATTTATCTATGATGAAACATCTTATGTAAACAGGAGCAAGTTGAT 1734
OY 1741 TCACCAAGAAAAAGTACCAACAAACAAAGGTAACCTTACCTTCAACAGGTGGCAAGGA 1800
DB 1735 TCTCTGAAAAAATATCCAAATTAACAACAAAGGTACATCTTCTTCAACAGGCGTAAGGA 1794
OY 1801 ATCTACGTTTACTTGAAGAGGCGCAATCTGCTACTATTTGCAAGAGTCTACTTGCT 1860
DB 1795 ATCTATGATATATGCGTGAGAGTATCTTCTTACTATGATGCTGACTGTTGCT 1854
OY 1861 AGACGTA 1867
DB 1855 AGACGCA 1861

RESULT 10
US-10-333-002-28
; Sequence 28, Application US/10333002
; Publication No. US20040071729A1
; GENERAL INFORMATION:
; APPLICANT: Adderson, Elisabeth
; APPLICANT: Bohnsack, John
; TITLE OF INVENTION: Group B Streptococcus Polypeptides Nucleic Acids and Therapeutic
; FILE REFERENCE: 2511-1-001 (Sf-0039)
; CURRENT APPLICATION NUMBER: US/10/333,002
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: PCT/US01/24795
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 09/534,341
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-10-333-002-28

Query Match 31.6%; Score 595; DB 8; Length 2199;
Best Local Similarity 100.0%; Pred. No. 2.2e-109;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1287 ACTAAATTCGTTACTTATGATTAATGAAAAAGCAAGTCCATGAAAGGCTTTA 1346
DB 1 ACTAAATTCGTTACTTATGATTAATGAAAAAGCAAGTCCATGAAAGGCTTTA 60
OY 1347 AAATGAGCGAATGTTGATTAACGTCATACGACGACCAACCAACCAACTGTTGAAGT 1406
DB 61 AAATGAGCGAATGTTGATTAACGTCATACGACGACCAACCAACCAACTGTTGAAGT 120
OY 1407 TGTGACAGGTGGAAAAGCTTTCATTAAGTCAATGCGCATGTGACACGACACACCTT 1466
DB 121 TGTGACAGGTGGAAAAGCTTTCATTAAGTCAATGCGCATGTGACACGACACACCTT 180
OY 1467 GGGGGAGCTTCCTTGTGTGCGTATGAAACAGCGACACAGCAATTAATTTGAAAT 1526
DB 181 GGGGGAGCTTCCTTGTGTGCGTATGAAACAGCGACACAGCAATTAATTTGAAAT 240
OY 1527 CGATGAAACAGCAAGCAACCTTGGTGAACCAAAAGCTGAACCACTACTTTAC 1586
DB 241 CGATGAAACAGCAAGCAACCTTGGTGAACCAAAAGCTGAACCACTACTTTAC 300
OY 1587 AACCAACGCTGATGATTTGTTGATATCACAGGGCTTTAAATAGCGTATTTATTAGA 1646
DB 301 AACCAACGCTGATGATTTGTTGATATCACAGGGCTTTAAATAGCGTATTTATTAGA 360
```

```
OY 1647 AGAACTGATGCTCTGATGATTAATGCTTTGTAACAAATGGAATTTGAGTCAA 1706
DB 361 AGAACTGATGCTCTGATGATTAATGCTTTGTAACAAATGGAATTTGAGTCAA 420
OY 1707 TGAACATCATATGAGCAACAAACAACTAGTTTCAACAGAAAAAGTACCAACAA 1766
DB 421 TGAACATCATATGAGCAACAAACAACTAGTTTCAACAGAAAAAGTACCAACAA 480
OY 1767 CAAAGTACCTTACTTCAACAGGTGGCAAAAGAAATCTACCTTCTTAGGAAGGCGC 1826
DB 481 CAAAGTACCTTACTTCAACAGGTGGCAAAAGAAATCTACCTTCTTAGGAAGGCGC 540
OY 1827 AGCTTGCTACTTATTTGAGAGTCTACTTGTCTAGAGTGAAGAAATGCT 1881
DB 541 AGCTTGCTACTTATTTGAGAGTCTACTTGTCTAGAGTGAAGAAATGCT 595

RESULT 11
US-10-425-115-120013/C
; Sequence 120013, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 120013
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1062)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: clone ID: MRT4577_40935C.1
US-10-425-115-120013

Query Match 3.6%; Score 67; DB 9; Length 1062;
Best Local Similarity 44.0%; Pred. No. 0.0037;
Matches 283; Conservative 0; Mismatches 360; Indels 0; Gaps 0;

OY 593 AAAAAGTGAACCTCTGAAATGAAGATTAAATGCGCGAGAAATTTGTTATTTCTAAA 652
DB 970 AAAAAGTGAACCTCTGAAATGAAGATTAAATGCGCGAGAAATTTGTTATTTCTAAA 911
OY 653 GCGAAGCTCACGAGCACGTAATATATTCAGAGAGTCAAAGATGATTAATATCACT 712
DB 910 AAAAAGTGAACCTCTGAAATGAAGATTAAATGCGCGAGAAATTTGTTATTTCTAAA 851
OY 713 GAGCAACGATTAACAAACAAACGCTTATTAATGCGGAAAGTTATGAAATTTGGG 772
DB 850 AAAAAGTGAACCTCTGAAATGAAGATTAAATGCGCGAGAAATTTGTTATTTCTAAA 791
OY 773 AAAATGATTTCAAGAGCAGAAATGGAACGCGAATTTAACATTAATTTCTGAGG 832
DB 790 AAAAAGTGAACCTCTGAAATGAAGATTAAATGCGCGAGAAATTTGTTATTTCTAAA 731
OY 833 TTGCTTCTGATTTTGAAGAGTGAAGCTCCAAATTAATGCAATTAATTTGAAATC 892
DB 730 AAAAAGTGAACCTCTGAAATGAAGATTAAATGCGCGAGAAATTTGTTATTTCTAAA 671
OY 893 AAAAAGTGAACCTCTGAAATGAAGATTAAATGCGCGAGAAATTTGTTATTTCTAAA 952
DB 670 AAAAAGTGAACCTCTGAAATGAAGATTAAATGCGCGAGAAATTTGTTATTTCTAAA 611
```

QY 953 AAAATGATACCTTAAGTGTGATTAACACACAGCTTAGGTGTAAGATGTGGCAA 1012
DB 610 AA 551
QY 1013 TTGGCGAAAAATTAATATCAATTTCTGTAAATATTCATTGGGATTCGACAAAG 1072
DB 550 AA 491
QY 1073 AAGGCGACCTAATTAATACGTCAATTCATTTAGTGAATCACTGATGACCTTAA 1132
DB 490 AA 431
QY 1133 CTTTGATTAACGCTTCTGAGAGATGATGCTTATATGATGAGGAGATACAGTGA 1192
DB 430 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 371
QY 1193 TTGCTCTGAAAAATTAATCAAGTGAACAAAGCAATGGCTT 1235
DB 370 ATATATATAGAACAAAAACAAAGCAAAAAACAAAGCGGCAT 328

RESULT 12

US-10-021-323-9336
; Sequence 9336, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, J111
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 9336
; LENGTH: 627
; TYPE: DNA
; ORGANISM: *Gossypium hirsutum*
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-024-Q6-K6-C2
US-10-021-323-9336

Query Match 3.5%; Score 65; DB 8; Length 627;
Best Local Similarity 46.1%; Pred. No. 0.0075;
Matches 218; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 649 AAAACCGAAGCTCACCGGCAAGTAAATATTCAGAGAGTCAAGATGATTTAT 708
DB 124 AAAAAAAAAAGGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 183
QY 709 ACATGACAAACGATTAAGAACAAACGCTTTATATCTGGGAAAGTATGAAAT 768
DB 184 AAAAAAAAAAAAAAAAAAAAAAAAAAGAGAAAAAAGAAAAAAGAAAAA 243
QY 769 GCGGAAATGATTTACAGAGACAGAAATGGAACGGGAATTAACAGTTAAATCTT 828
DB 244 GGAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 303
QY 829 GAGGTGCTGCTATATTTTGAAGAGTAAAGCTCCAAATTAATGCAATTAATTGA 888
DB 304 AAGGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 363
QY 889 AATCAACAAAAACACATTTTCAATTGAGCAAAACATCAACACTGTGAAAAACA 948
DB 364 AA 423
QY 949 GTCAAAATATGATCTTAAGTGAATTAACACACCAAGCTTGAAGTGAAGATGTG 1008
DB 424 AAAAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 483

QY 1009 GCAATTTGGGAAAAATTAATATCAATTTCTGTAATATTCATTTGGGATTCAGAC 1068
DB 484 AA 543
QY 1069 AAGAGGCGACGCTAATTAATACGTCAATTTCAATTTAGTGAATCAATGTA 1121
DB 544 AAAAAAAAAAGAAACCAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 596

RESULT 13

US-10-172-502-1
; Sequence 1, Application US/10172502
; Publication No. US20030185833A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.
; FILE REFERENCE: P07263US01/BAS
; CURRENT APPLICATION NUMBER: US/10/172,502
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6609
; TYPE: DNA
; ORGANISM: *Staphylococcus epidermidis*
US-10-172-502-1

Query Match 3.4%; Score 64.6; DB 7; Length 6609;
Best Local Similarity 46.9%; Pred. No. 0.023;
Matches 202; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 545 ATATTTATCTTAATTAATGATGAGCAATGATGATGATTTACATGTAAGAAAAAGTAGAA 604
DB 544 ATACGAGGCGACAGATGAAGAAACAAATGCTGCAATAGTACAGTTGAAAAAGATTAA 5608
QY 605 CTGCTGAAAAATGAAGATTTAAATGGCGCAATTTGTTATTTCTAAGAACGAGCTCAG 664
DB 5609 TTAAAGCTTAACAAATATGCTGAGTGAAGTCAATGCTGATGCTATTTATTTATGTC 5668
QY 665 CAGGACAGTAAATTAATTCAGAGAGTCAAGATGATTAATATCATGACACAGGATA 724
DB 5669 ATGATGGAAGAAACAAATTTGCTGAATCGAACTGTTATTAATTAAGAACTGCGCC 5728
QY 725 AAGAACAGCAAAACGCTTTATTTACTGGAAGAAAGTTAGAAATGGCGAAATGATTGA 784
DB 5729 GAGAACATTAACACATTTATTAACGATTAAGAAACAGCAATTAAGGAAATGTTCAAG 5788
QY 785 CAGAACGAGAAATGAACGGAGAAATTAACAGTTAAATTTGAGTTGTTGTTGTTATA 844
DB 5789 CAACGATGAAGAAAGAAATGATTTATTTAGCAGTTCAAAACATTAATGACACTGCTA 5848
QY 845 TTTTGAAGAAATTAAGCTCAATTAATGCAAGATTAATGAATCAACAAACAC 904
DB 5849 TTGACAAATGATTAAGATGCTAGCAATGCAAGATGATTAACACAGCACTTAATTC 5908
QY 905 CATTTACATTAAGCAAAACATCAACACCTGTTGAAGAAAAACAGTCAAAAAATGATCCT 964
DB 5909 TACAAACATATCATGATTTAGACGTACATCTTAATTAAGCAAGCATGCTGAAAAACGA 5968
QY 965 CTAAAGTTGAT 975
DB 5969 TTAATGATGAT 5979

RESULT 14

US-11-020-509-1
; Sequence 1, Application US/11020509
; Publication No. US20050106648A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.

```

; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.
; FILE REFERENCE: P07263US02/BAS
; CURRENT APPLICATION NUMBER: US/11/020,509
; CURRENT FILING DATE: 2004-12-27
; PRIOR APPLICATION NUMBER: US 10/172,502
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6609
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; US-11-020-509-1

```

```

; NAME/KEY: unsure
; LOCATION: (1)..(1029)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_43736C.1
US-10-425-115-123086

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 6, 2006, 00:20:38 ; Search time 176 Seconds
(without alignment) 13895.170 Million cell updates/sec

Title: US-10-661-809a-12
Perfect score: 1881
Sequence: 1 atgaagcaatcaaaaaagc.....gacgtagaagaagaatgct 1881

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 844069 seqs, 65006433 residues

Total number of hits satisfying chosen parameters: 1668138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA New:
1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB_seq.*
2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB_seq.*
3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB_seq.*
4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB_seq.*
5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCF_NEW_PUB_seq.*
6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB_seq.*
7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB_seq.*
8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB_seq.*
9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.8	3.2	2596	US-11-218-305-12130	Sequence 12130, A
2	56.6	3.0	6558	US-10-471-571A-859	Sequence 859, App
3	54.6	2.9	4009	US-11-218-305-3660	Sequence 3660, App
4	54.2	2.9	2522	US-11-218-305-24510	Sequence 24510, A
5	52	2.8	2718	US-11-218-305-12773	Sequence 12773, A
6	51.4	2.7	1826	US-11-218-545-774	Sequence 774, App
7	51.4	2.7	5493	US-10-517-441-691	Sequence 691, App
8	50	2.7	1407	US-10-471-571A-2917	Sequence 2917, App
9	49.8	2.6	19634	US-10-517-441-413	Sequence 413, App
10	49.8	2.6	19634	US-10-517-441-687	Sequence 687, App
11	49.6	2.6	951	US-11-266-748A-48160	Sequence 48160, A
12	49.6	2.6	5493	US-10-517-441-417	Sequence 417, App
13	49.2	2.6	1000	US-11-266-748A-222640	Sequence 222640, A
14	49.2	2.6	1000	US-11-266-748A-289165	Sequence 289165, A
15	49.2	2.6	1000	US-11-266-748A-340594	Sequence 340594, A
16	49.2	2.6	1000	US-11-266-748A-400143	Sequence 400143, A
17	49.2	2.6	1000	US-11-266-748A-471189	Sequence 471189, A
18	49	2.6	8759	US-10-517-441-530	Sequence 530, App
19	48.8	2.6	10865	US-10-517-441-463	Sequence 463, App
20	48.8	2.6	49143	US-11-021-837-34	Sequence 28366, App
21	48.6	2.6	4619	US-11-266-748A-28366	Sequence 354, App
22	48.2	2.6	4930	US-10-517-441-354	Sequence 628, App
23	48.2	2.6	4930	US-10-517-441-628	Sequence 511, App
24	48.2	2.6	6001	US-10-517-441-511	Sequence 511, App

C 25	48.2	2.6	6001	6	US-10-517-441-785	Sequence 785, App
C 26	48.2	2.6	7833	6	US-10-517-441-350	Sequence 350, App
C 27	48.2	2.6	7833	6	US-10-517-441-624	Sequence 624, App
C 28	48	2.6	6521	6	US-10-517-441-584	Sequence 584, App
C 29	47.4	2.5	1147	7	US-11-218-305-2931	Sequence 2931, App
C 30	47.4	2.5	8759	6	US-10-517-441-256	Sequence 256, App
C 31	47.2	2.5	6521	6	US-10-517-441-310	Sequence 310, App
C 32	47	2.5	8666	6	US-10-517-441-558	Sequence 558, App
C 33	47	2.5	1237661	8	US-11-266-748A-29041	Sequence 29041, A
C 34	46.8	2.5	7434	6	US-10-471-571A-2277	Sequence 2277, App
C 35	46.4	2.5	2025	6	US-10-531-659-8	Sequence 8, App
C 36	46.4	2.5	2412	8	US-11-217-529-6042	Sequence 6042, App
C 37	46	2.4	4047	6	US-10-471-571A-3351	Sequence 3351, App
C 38	46	2.4	163280	7	US-11-021-837-47	Sequence 47, App
C 39	45.8	2.4	3223	6	US-10-517-441-388	Sequence 388, App
C 40	45.8	2.4	3223	6	US-10-517-441-682	Sequence 682, App
C 41	45.8	2.4	28605	6	US-10-471-571A-4495	Sequence 4495, App
C 42	45.6	2.4	1000	8	US-11-266-748A-293117	Sequence 293117, A
C 43	45.6	2.4	1000	8	US-11-266-748A-344546	Sequence 344546, A
C 44	45.6	2.4	1000	8	US-11-266-748A-405014	Sequence 405014, A
C 45	45.6	2.4	1000	8	US-11-266-748A-476060	Sequence 476060, A

ALIGNMENTS

RESULT 1									
US-11-218-305-12130									
Sequence 12130, Application US/11218305									
Publication No. US2006014195A1									
GENERAL INFORMATION:									
APPLICANT: MONSANTO TECHNOLOGY, LLC									
APPLICANT: McLaif, Paul L.									
APPLICANT: Tao, Nengsheng									
APPLICANT: Wu, Kunsheng									
TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping									
FILE REFERENCE: 38-21 (53660)B									
CURRENT FILING DATE: 2005-09-01									
PRIOR APPLICATION NUMBER: US 60/606,880									
NUMBER OF SEQ ID NOS: 25043									
SOFTWARE: Patent version 3.2									
SEQ ID NO 12130									
LENGTH: 2596									
TYPE: DNA									
ORGANISM: Zea mays									
US-11-218-305-12130									
Query Match									
Best Local Similarity 44.1%; Pred. No. 0.00014;									
Matches 250; Conservative 0; Mismatches 317; Indels 0; Gaps 0;									
QY	552	TCCTAAATGTCGTAGCCATGATGATGTTTACATGCAAAAGTAGAAGTCTGCA	611						
DB	1823	TCTTATGCTAGGAGAGAGACGACTGCTGTTACAGAGAAAAA	1882						
QY	612	AAATGAGATTAATGCGGAGATTTGTTATTTCTAAAGCGAAGCTCACGCGC	671						
DB	1883	AAAAAAGATTAATGCGGAGATTTGTTATTTCTAAAGCGAAGCTCACGCGC	1942						
QY	672	ACTTAAATATTCACAGAGTCGAAAGATTAATACATGCAAGCGATTAAGACA	731						
DB	1943	AAAAAAGATTAATTCACAGAGTCGAAAGATTAATACATGCAAGCGATTAAGACA	2002						
QY	732	AGCAAAAGCTTTTATTAAGGAAAGTTATGCGCAAAATGATTTTACAGAGC	791						
DB	2003	AAAAAAGCTTTTATTAAGGAAAGTTATGCGCAAAATGATTTTACAGAGC	2062						
QY	792	AGAGATGAGACGGAGATTAACGATTAATAATCTTGAGGTGCTGATATTTTGA	851						
DB	2063	AAAAAAGCTTTTATTAAGGAAAGTTATGCGCAAAATGATTTTACAGAGC	2122						


```
Qy 852 AGAAGTAAAGCTCCAAATATGCAATTAATTGAAATCAAGAAAAACCATTTAC 911
Db 2123 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2182
Qy 912 AATTGAGCAAAACAATCAACACCTGTTGAAAAAAGCTCAAAATGATCCTTAAGT 971
Db 2183 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2242
Qy 972 TGATTAACACACCAAGCTTAGATGATGATGATGATGATGATGATGATGATGAT 1031
Db 2243 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2302
Qy 1032 TCAATTTCTGTAATATTTCCATTGGGAGTTGCAGACAAAGAGCGCATTAATA 1091
Db 2303 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2362
Qy 1092 CGTCAAATTCATTTAGTTGATTAACA 1118
Db 2363 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2389
```

```
RESULT 2
US-10-471-571A-859
; Sequence 859, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqMan9, version 1.03
; SEQ ID NO 859
; LENGTH: 6558
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-859
```

```
Query Match 3.0%; Score 56.6; DB 6; Length 6558;
Best Local Similarity 45.7%; Pred. No. 0.0012;
Matches 197; Conservative 0; Mismatches 234; Indels 0; Gaps 0;
Qy 545 ATATTATCTTAAATGTGATGCAATGATGATGATGATGATGATGATGATGATGAT 604
Db 5501 ATAAAGAAAGCGACAGATGAAGAAATAATGCAATGACACAAAGTTGAAAAAGATTAA 5560
Qy 605 CTGCTGAAAAATGAAGATTAATGCGGAGATTTGTTATTTCTAAAAGCGAAGCTCAC 664
Db 5561 TTAAAGCTTAAACAATAATGCTAGTGACGATGATGATGATGATGATGATGATGATGAT 5620
Qy 665 CAGGACAGTAAATATATCCAAAGATCAAAAGATGATTAATATATATATATATATATAT 724
Db 5621 ATGATGAAAAAAGCAAAATGCTGAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 5680
Qy 725 AAGAAACAAGCAAAAGCTTTATTTACTGGGAAAAAGTTAAGAAATGGCGAAAAATGATTCA 784
Db 5681 GAGAACAATTGCAACATTAATTCACAGATTAATAAAGCAATTAAGGAATATTTCAG 5740
Qy 785 CAGAAAGAGAAATGAAACGGGAAATTAACGTTAAATTTGAGGTTGGTGTGATA 844
Db 5741 CAACGTTAGAAAGAAATATATATATATATATATATATATATATATATATATATATATAT 5800
Qy 845 TTTTGAAGAAATGAAAGCTCCAAATTAATGCAATTAATGCAATTAATGCAATTAATGCAAT 904
Db 5801 TTGCAAAATGATCAAGATGCTAGCAATGCAAAAGTTGATTAATAAGATCATTTAAATC 5860
Qy 905 CATTTTCAATTGAGCAAAACAATCAAAACCTGTTGAAAAAAGCTCAAAATGATACCT 964
Db 5861 TACAAACAATACATGATTTAGATTTGATTCATTTTAAAGCCAGATGCTGAAAAAAGCA 5920
```

```
Qy 965 CTAAAGTTGAT 975
Db 5921 TTAATGATGAT 5931
```

```
RESULT 3
US-11-218-305-3660
; Sequence 3660, Application US/11218305
; Publication No. US2006014195A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McIaird, Paul L.
; APPLICANT: Tao, Nenpheng
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: Corn.
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3660
; LENGTH: 4009
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (22)..(22)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (111)..(111)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (178)..(178)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (235)..(235)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (285)..(285)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (290)..(290)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (308)..(308)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (310)..(313)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (341)..(341)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (362)..(364)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (408)..(408)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-3660
```


Query Match 2.9%; Score 54.6; DB 7; Length 4009;
Best Local Similarity 44.8%; Pred. No. 0.003; Indels 0; Gaps 0;
Matches 174; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

Qy 649 AAAGCGAGGCTCACGAGCAGTAAATATATATCAAGATCAAGATGATATAT 708
Db 3 AAAAAAAAAAGAGCGGNNAAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAGAGAG 62

Qy 709 ACATGACACCGATTAAGACAGCAAAACCTTATATCTGGGAAAAGTTATGAATT 768
Db 63 TAAAAAAGACAGCAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 122

Qy 769 GCGCAAAATGATTTACAGAGCAGAGATGGAACGGGAGATTAAACATTAAATCTT 828
Db 123 GGGGAAAG 182

Qy 829 GAGTTGTTCTATATTTTGAAGAGTAAAGCTCAATATATGCAAGATTATGAA 888
Db 183 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 242

Qy 889 AATCAACAAAAACACCATTTACAATTGAGCAACATCAACCTGTTGAAAAACA 948
Db 243 AAAAAAAAAAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 302

Qy 949 GTCAAAATGATCTTAAGTTAAGCAACCAAGCTTATGATGTAAGATGTC 1008
Db 303 AAAAAAAAAAAAAAAAAAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 362

Qy 1009 GCAATTGGCGAAAAATTAATATCAAA 1036
Db 363 NNGAAGGAGAGAAAAAGAAAAAGAA 390

RESULT 4
US-11-218-305-24510/c
; Sequence 24510, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaoid, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Corn
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218.305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: Patencin version 3.2
; SEQ ID NO 24510
; LENGTH: 2522
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-24510

Query Match 2.9%; Score 54.2; DB 7; Length 2522;
Best Local Similarity 45.1%; Pred. No. 0.0031;
Matches 200; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

Qy 594 AAAAGTAGAAGCTGTGAAAAAGAGATTAATGCGGAGATTGTTATTTCTAAAG 653
Db 447 AAAAAATTAATGATCTTAACAACTTAATATATATATGTTATTAACGGAATGTAC 388

Qy 654 CGAAGCTCACGAGCAGCAAGTAAATATATCAAGAGTCAAGATGATATACATG 713
Db 387 CGTTTCAGTAG 328

Qy 714 GACAAAGATTAAGAAACAAAGCAACGCTTTATTACTGGGAAAGTTATGAATGGCGA 773
Db 327 AA 268

Qy 774 AAATGATTTTCAGAGAGCAGAGATGAACGGAGATTATACAGTTAAATCTTGAGGT 833
Db 267 AA 208

Qy 834 TGGTTCTATATTTTGAAGAGTAAAGCTCAATATATGCAATTAATTAATGAATGA 893
Db 207 AA 148

Qy 894 AACAAAAACCATTTTACATTTGAAGCAACATCAACCTGTTGAAAAACGTCAA 953
Db 147 AA 88

Qy 954 AAATGATCTCTAAGTTATTAAGTAAACACCAAGCTTATGATGTTAAAGATGGCAAT 1013
Db 87 AA 28

Qy 1014 TGGCGAAAAATTAATATCAAA 1036
Db 27 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5

RESULT 5
US-11-218-305-12773/c
; Sequence 12773, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaoid, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Corn
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218.305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: Patencin version 3.2
; SEQ ID NO 12773
; LENGTH: 2718
; TYPE: DNA
; ORGANISM: Zea mays

NAME/KEY: misc feature
LOCATION: (2410)..(2410)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (2425)..(2430)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (2434)..(2434)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (2442)..(2442)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (2504)..(2504)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (2510)..(2510)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (2570)..(2570)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:

```
NAME/KEY: misc_feature
LOCATION: (2595)..(2599)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2613)..(2613)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2633)..(2633)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2675)..(2675)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2680)..(2680)
OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-12773
```

```
Query Match 2.8%; Score 52; DB 7; Length 2718;
Best Local Similarity 44.3%; Pred.No. 0.011;
Matches 181; Conservative 0; Mismatches 228; Indels 0; Gaps 0;
```

```
QY 585 ACATGTGAAAAAGTAGGAACTGCTGAAAAATGAGGATTAATGGCGCAGATTGTTAT 644
DB 2681 ANAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2622
QY 645 TTCCTAAAGCGAGGCTCCAGGCAAGTAAATATATCCAGAGATCAAGATGAT 704
DB 2621 ATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2562
QY 705 ATATCATGAGCAAGGATTAAGCAAGCAAGCTTTATCTGGGAAAGTTATGA 764
DB 2561 GAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2502
QY 765 AATTGCGCAAAATGATTTCAAGAGAGAGATGAAACGAGATTTACAGTTAAAA 824
DB 2501 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2442
QY 825 TCTTGAGGTGCTGCTATATTTTGAAGAGATTAAGCTCAATTAATGCAATTAT 884
DB 2441 GAGAGAGNAAANNAATTAATAAATAAATAAATAAATAAATAAATAAATAA 2382
QY 885 TGAATAATGAACAAACACCATTTACATTTGAAGCAATCAACACCTGTGAAA 944
DB 2381 AAAAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2322
QY 945 AACAGTCAAAATGATACCTTAAGTTGATTAACCAACACCAAGCTTA 993
DB 2321 GAAAAAGAGAAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2273
```

```
RESULT 6
US-11-216-545-774
Sequence 774, Application US/11216545
Publication No. US20060135758A1
GENERAL INFORMATION:
APPLICANT: MONSANTO Technology, LLC
APPLICANT: McLaoid, Paul L
APPLICANT: Mu, Kunsheng
APPLICANT: Tao, Nengping
TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
FILE REFERENCE: 38-21 (53659)B
CURRENT APPLICATION NUMBER: US/11/216,545
PRIOR APPLICATION NUMBER: 2005-08-31
PRIOR FILING DATE: 2004-08-31
NUMBER OF SEQ ID NOS: 8783
SOFTWARE: PatentIn version 3.2
SEQ ID NO 774
```

```
LENGTH: 1826
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10)..(10)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15)..(16)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (31)..(31)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (33)..(39)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (42)..(43)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45)..(46)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51)..(53)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (58)..(61)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (73)..(73)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (75)..(80)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (97)..(97)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100)..(103)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (106)..(109)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (112)..(113)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (118)..(121)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (125)..(125)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
```

LOCATION: (136)..(141)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (148)..(152)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (156)..(158)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (171)..(172)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (175)..(175)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (190)..(192)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (196)..(198)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (214)..(218)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (224)..(224)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (226)..(228)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (230)..(232)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (247)..(248)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (252)..(252)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (254)..(255)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (263)..(264)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (269)..(273)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (275)..(277)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (283)..(285)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (288)..(288)

OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (291)..(293)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (297)..(297)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (300)..(301)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (308)..(310)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (313)..(317)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (322)..(324)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (331)..(331)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (333)..(335)
OTHER INFORMATION: n 1s a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (340)..(340)
OTHER INFORMATION: n 1s a, c, g, or t

Query Match 2.7%; Score 51.4; DB 8; Length 1826;
Best Local Similarity 29.4%; Pred. No. 0.013;
Matches 127; Conservative 0; Mismatches 305; Indels 0; Gaps 0;

QY 670 ACAGTAAATATATTCACAGAGTCAGATGATTTATATCATGACAAAGGATTAAGAA 729
DB 20 AA 79
QY 730 CAAGCAAAACGCTTATTACTGGAAAGTTATGAATTGGCGAAATGATTTACAGAA 789
DB 80 NAAA 139
QY 790 GCAGAGATGGAACGGAGATTAACAGTTAAAAATCTTGAGGTTGCTGATATTTA 849
DB 140 NAAA 199
QY 850 GAAGAGTAAAGCTCAATATATGAGAAATTAATGAAATCAACAAACACACATTT 909
DB 200 AA 259
QY 910 ACAATTTGAAGCAACATCAACACCTGTTGAAAAACGTCAAAAATGATACCTTAA 969
DB 260 AA 319
QY 970 GTTGATTAACACACACCTGTTAGTGTAAAGATGTGCAATTGGCGAAAAATTTAA 1029
DB 320 AA 379
QY 1030 TATCAATTTCTGTAATATTCATGCGGATTCAGACAAAGCGACGCTAATAA 1089
DB 380 AA 439
QY 1090 TACGTCAAAATTC 1101
DB 440 AGAGTCAGCTAC 451

```
RESULT 7
US-10-517-441-691/c
; Sequence 691, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: POEKENS, John
; APPLICANT: HARBEC, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARK, Almut
; APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; TITLE OF INVENTION: Proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 691
; LENGTH: 5493
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-691

Query Match      2.7%; Score 51.4; DB 6; Length 5493;
Best Local Similarity 46.1%; Pred. No. 0.019;
Matches 172; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 676 AATATATCCAGAGTCAAGATGATTATATACATGACCAACGATNAGACAGCA 735
DB 5477 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAT 5418
QY 736 AAAGCTTATTTACTGGGAAAAGTTATGAAATTTGGGAAAATGATTTTCACAGACGAG 795
DB 5417 AAAAAAAAACAACAAAAAATATCAACAAAAAACAACCAAAATTAATAAATAA 5358
QY 796 AATGACGGAGATTAACAGTAAATCTTGAGTTGTTGATTTTGAAGA 855
DB 5357 AACAATCAATTAATTAATAAACAACAAAAAACAACAAAAAACAACAAAAAATAA 5298
QY 856 GTAAAGCTCCAAATATGACAGATTAATTTGAAAAATCAACAAAAAACAACATTTCAATT 915
DB 5297 AAAAAAAAACAACAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 5238
QY 916 GAGCAAAACATCAACACCTGTTGAAAAAACAATGATTAACCTTAAGTTGAT 975
DB 5237 AAAAAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5178
QY 976 AAAAAACAACAAGCTTGATGATTAAGTGGCAATTTGGGAAAAATTAATATCA 1035
DB 5177 CAATAAACAATAAATAAACAACAACAATAAATAAATAAATAAATAAATAAATAA 5118
QY 1036 ATTTCTGTAATA 1048
DB 5117 ACACCAAAAAA 5105
```

```
RESULT 8
US-10-471-571A-2917
; Sequence 2917, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 2917
; LENGTH: 1407
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-471-571A-2917

Query Match      2.7%; Score 50; DB 6; Length 1407;
Best Local Similarity 45.6%; Pred. No. 0.026;
Matches 176; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 754 AAAAGTTATGAATGGGAAAATGATTTTCACAGACGAGATGAAACGGAGAAATTA 813
DB 322 AAAAAAGATTAACCCGAAATTTGAAAGATTTATGAGAGAGCAATTAAGTGCACCTTA 381
QY 814 ACAGTTAAATCTTGAGTTGTTGATTAATTTTGAAGAAGTAAAGCTCAATAT 873
DB 382 GAATTAACAATTAAGAAATCAAGATTAATTTGTTAGTAAACCTTTTATCAAACTAT 441
QY 874 GCAGATTAATTAATCAACAAACAAACACCAATTTACATTTGAACAAACATCAACA 933
DB 442 AGAGATGATGTTGAAGATTTATATAGTAAGTATTAATTAATTAATGAGATTAAGATGA 501
QY 934 CCTGTGAAAAACAGTCAAAATATATCTTAAAGTTGATTAACAAACACCAAGCTTA 993
DB 502 GAAAGAGCAATTAATAAAGCAAGTTAACAAGATTTGAAATTAATAAATAAATAAATAA 561
QY 994 GATGTAAGATGTCGCAATTTGCGCAAAATTAATATCAATTTCTGTAATATTTCA 1053
DB 562 GAAACATTAATTAATTAATTTTGTGATTAATTAATTAATTAATTAATTAATTTCT 621
QY 1054 TTGGGATTTGACAGCAAAAGGCGCGCTAATTAATTAATTAATTAATTAATTAATTA 1113
DB 622 GTTTAAGAAATGAATAAACAAGAAAGAAATTAATAAATAAATGCTCAATTAATTAATCT 681
QY 1114 AAACATGATGACGCTTAACCTTTGA 1139
DB 682 GACACTGAAGCAGCAAAAGATGA 707

RESULT 9
US-10-517-441-413/c
; Sequence 413, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: POEKENS, John
; APPLICANT: HARBEC, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARK, Almut
; APPLICANT: HOEFER, Heinz
```

```

; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 413
; LENGTH: 19634
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-413
```

```

Query Match      2.6%; Score 49.8; DB 6; Length 19634;
Best Local Similarity 46.7%; Pred. No. 0.073;
Matches 192; Conservative 0; Mismatches 217; Indels 2; Gaps 1;
```

```

QY 705 ATATCATGACGACGATTAAGAACCAAGCCTTTATTCTGGGAAAGTTATGA 764
   |||||
DB 6994 AAAAACTAATATAAAAAAATTAACCAAAAAAATTAATTAATTAATTA 6935
   |||||
QY 765 AATTGGCGAAATGATTTCAAGAGCAGAGATGGAATGGAATTAACAGTTAA 824
   |||||
DB 6934 AATT--CAATCATATCTATTAAAAATTAATTAATTAATTAATTAATTA 6877
   |||||
QY 825 TCTTAGGTTGTTGCTATATTTTGAAGAGTAAAGCTCCAAATATGAGAAATTA 884
   |||||
DB 6876 TAAAAACACTAAAAAATTAACATCTAATAAAAAATTAATTAATTAATTA 6817
   |||||
QY 885 TGAATATCAAAACAAACACCTTTTCAATTTGAGCAACATCAACACCTGTGAAA 944
   |||||
DB 6816 CTATTAATAAAATATTAATCAATTAATTAATTAATTAATTAATTAATTA 6757
   |||||
QY 945 AACAGTCAAAAATGATACCTTAAGTTGATTAACCAACCAAGCTTATGATGTA 1004
   |||||
DB 6756 AATTAACAAATTAATAAATTAACAAATTTTAATAATTAACAAATTAATCA 6697
   |||||
QY 1005 TGTGCAATTGGCGAAAAATTAATTAATCAATTTCTGTAATTAATTCATTGG 1064
   |||||
DB 6696 ATTTTAAATAATCTTAATAATCAAAATTTTAATAAATAATTAATAAATA 6637
   |||||
QY 1065 AGACAAAGAGCGACGCTAATTAATTAATTAATTAATTAATTAATTAATTA 1115
   |||||
DB 6636 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6586
   |||||
```

RESULT 10
US-10-517-441-687/c

; Sequence 687, Application US/10517441
; Publication No. US20060121467a1

; GENERAL INFORMATION:

; APPLICANT: HARBEC, John
; APPLICANT: FOKKENS, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almut
; APPLICANT: HOEBLER, Heinz

```

; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 687
; LENGTH: 19634
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-687
```

```

Query Match      2.6%; Score 49.8; DB 6; Length 19634;
Best Local Similarity 46.7%; Pred. No. 0.073;
Matches 192; Conservative 0; Mismatches 217; Indels 2; Gaps 1;
```

```

QY 705 ATATCATGACGACGATTAAGAACCAAGCCTTTATTCTGGGAAAGTTATGA 764
   |||||
DB 6994 AAAAACTAATATAAAAAAATTAACCAAAAAAATTAATTAATTAATTA 6935
   |||||
QY 765 AATTGGCGAAATGATTTCAAGAGCAGAGATGGAATGGAATTAACAGTTAA 824
   |||||
DB 6934 AATT--CAATCATATCTATTAAAAATTAATTAATTAATTAATTAATTA 6877
   |||||
QY 825 TCTTAGGTTGTTGCTATATTTTGAAGAGTAAAGCTCCAAATATGAGAAATTA 884
   |||||
DB 6876 TAAAAACACTAAAAAATTAACATCTAATAAAAAATTAATTAATTAATTA 6817
   |||||
QY 885 TGAATATCAAAACAAACACCTTTTCAATTTGAGCAACATCAACACCTGTGAAA 944
   |||||
DB 6816 CTATTAATAAAATATTAATCAATTAATTAATTAATTAATTAATTAATTA 6757
   |||||
QY 945 AACAGTCAAAAATGATACCTTAAGTTGATTAACCAACCAAGCTTATGATGTA 1004
   |||||
DB 6756 AATTAACAAATTAATAAATTAACAAATTTTAATAATTAACAAATTAATCA 6697
   |||||
QY 1005 TGTGCAATTGGCGAAAAATTAATTAATCAATTTCTGTAATTAATTCATTGG 1064
   |||||
DB 6696 ATTTTAAATAATCTTAATAATCAAAATTTTAATAAATAATTAATAAATA 6637
   |||||
QY 1065 AGACAAAGAGCGACGCTAATTAATTAATTAATTAATTAATTAATTAATTA 1115
   |||||
DB 6636 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6586
   |||||
```

RESULT 11
US-11-266-748a-48160

; Sequence 48160, Application US/11266748A
; Publication No. US2006013463a1

; GENERAL INFORMATION:

; APPLICANT: Hartin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4

PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105507.0
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105485.9
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: EP 04105484.2
PRIOR FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: US 60/662,276
PRIOR FILING DATE: 2005-03-14
PRIOR APPLICATION NUMBER: US 60/700,293
PRIOR FILING DATE: 2005-07-18
NUMBER OF SEQ ID NOS: 483996
SOFTWARE: PatentIn version 3.3
SEQ ID NO 48160
LENGTH: 951
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (528)..(528)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (558)..(558)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (560)..(560)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (562)..(562)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (564)..(564)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (566)..(566)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (577)..(579)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (581)..(583)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (585)..(590)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (592)..(593)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (605)..(611)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (614)..(619)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (623)..(625)
OTHER INFORMATION: n is a, c, g, or t

FEATURE:
NAME/KEY: misc_feature
LOCATION: (630)..(631)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (635)..(636)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (638)..(639)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (641)..(641)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (656)..(661)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (666)..(675)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (685)..(685)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (687)..(698)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (701)..(701)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (705)..(705)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (707)..(707)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (712)..(712)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (715)..(716)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (725)..(725)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (732)..(733)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (738)..(739)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (741)..(743)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (746)..(746)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:

```

NAME/KEY: misc feature
LOCATION: (761)..(762)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (768)..(769)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (771)..(771)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (773)..(773)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (782)..(782)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (800)..(800)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (802)..(805)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (821)..(824)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (826)..(831)

```

```

Query Match      2.6%; Score 49.6; DB 8; Length 951;
Best Local Similarity 30.8%; Pred. No. 0.028;
Matches 121; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

```

```

Qy 612 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 671
Db 559 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 618
Qy 672 AGTAATAATATCCAGGAGTCAAGATGATTATATACAGCAACGATTAAGAACA 731
Db 619 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 678
Qy 732 AGCAAAAGCTTATTAATCTGCGAAAGTTATGAAATGCGGAAATGATTTTCAAGAAC 791
Db 679 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 738
Qy 792 AGAAGATGAGCGGAGATTTAAACGTTAAATCTTGAAGTTGCTGATTTTGA 851
Db 739 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 798
Qy 852 AGAAGTAAAGCTCAAGATTAATGAGATTTAATGAAATCAACAAACACCACTTAC 911
Db 799 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 858
Qy 912 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 971
Db 859 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 918
Qy 972 TGATAAACAACACGAGCTTAGATGTTAAGA 1004
Db 919 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 951

```

RESULT 12

```

US-10-517-441-417/c
Sequence 417, Application US/10517441
Publication No. US20060121467A1
GENERAL INFORMATION:
APPLICANT: FOKKENS, John
APPLICANT: HARBESCK, Nadia
APPLICANT: KOENIG, Thomas
APPLICANT: MAIER, Sabine
APPLICANT: MARTENS, John
APPLICANT: MODELL, Fabian
APPLICANT: NIMMERICH, Inko
APPLICANT: RUJAN, Tamas
APPLICANT: SCHMITT, Armin
APPLICANT: SCHMITT, Manfred
APPLICANT: LOOK, Maxime P.
APPLICANT: MARX, Almut
APPLICANT: HOEPLER, Heinz
TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
FILE REFERENCE: 47675-93
CURRENT APPLICATION NUMBER: US/10/517,441
CURRENT FILING DATE: 2004-12-11
PRIORITY APPLICATION NUMBER: PCT/EP2003/010881
PRIORITY FILING DATE: 2003-10-01
PRIORITY APPLICATION NUMBER: DE 10317955.0
PRIORITY FILING DATE: 2003-04-17
PRIORITY APPLICATION NUMBER: DE 10300096.8
PRIORITY FILING DATE: 2003-01-07
PRIORITY APPLICATION NUMBER: DE 10245779.4
PRIORITY FILING DATE: 2002-10-01
NUMBER OF SEQ ID NOS: 2147
SEQ ID NO 417
LENGTH: 5493
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-417

```

```

Query Match      2.6%; Score 49.6; DB 6; Length 5493;
Best Local Similarity 44.3%; Pred. No. 0.052;
Matches 202; Conservative 0; Mismatches 254; Indels 0; Gaps 0;

```

```

Qy 648 TAAAGCGAGCTCAACGAGTAAATATTCAGAGATCAAGATGATTA 707
Db 5370 TAAAGCGAGCTCAACGAGTAAATATTCAGAGATCAAGATGATTA 5311
Qy 708 TACATGACCAACGAGTAAAGCAAGCAACGCTTTATTAATGAGTAAT 767
Db 5310 TAAAGCGAGCTCAACGAGTAAATATTCAGAGATCAAGATGATTA 5251
Qy 768 TGGCGAATATGATTTTCAAGAGAGAGATGAGCGGAGATTAACGTTAAATCT 827
Db 5250 TAAAGCGAGCTCAACGAGTAAATATTCAGAGATCAAGATGATTA 5191
Qy 828 TGAGTTGCTGATTAATTTTGAAGATTAATCAATTAATGCAATTAATGA 887
Db 5190 TAAAGCGAGCTCAACGAGTAAATATTCAGAGATCAAGATGATTA 5131
Qy 888 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 947
Db 5130 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 5071
Qy 948 AGTCAAAATGATCTCTTAAGTTGATTAACCAACCAAGCTTAGATGTTAAGATGT 1007
Db 5070 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 5011
Qy 1008 GGCATTTGGCGAATTAATTAATCAATTTCTGTAATTTTCATTTGGGATTCAGA 1067
Db 5010 AATTAAGATTAATGCGCGAATTTCTTAAAGCGAGCTCACGCGAC 4951
Qy 1068 CAAAGAGCGAGCTTAATTAATGTTCAATTTCA 1103

```



```

? PRIOR APPLICATION NUMBER: US 60/700,293
?
? PATENT FILING DATE: 2005-07-18
?
? NUMBER OF SEQ ID NOS: 483,996
?
? SOFTWARE: PatentIn version 3.3
?
? SEQ ID NO 340594
?
? LENGTH: 1000
?
? TYPE: DNA
?
? ORGANISM: Homo Sapiens
?
? US-11-266-748A-340594

```

Query Match	2.6%	Score 49.2;	DB 8;	Length 1000;
Best Local Similarity	50.9%;	Pred. No. 0.036;		
Matches 117; Conservative	0;	Mismatches 113;	Indels 0;	Gaps 0;

[illegible]

Search completed: July 6, 2006, 00:51:23
Job time : 177 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2006, 20:23:58 ; Search time 265 Seconds
(without alignments)
13281.342 Million cell updates/sec

Title: US-10-661-809a-12

Perfect score: 1881

Sequence: 1 atgaagcaatttaaaagct.....gacgtgagaagaagaatgct 1881

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:
1: /EMC Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:
2: /EMC Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:
3: /EMC Celerra_SIDS3/ptodata/2/ina/6 COMB.seq:
4: /EMC Celerra_SIDS3/ptodata/2/ina/6 COMB.seq:
5: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:
6: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:
7: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:
8: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:
9: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:
10: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1881	100.0	1980	US-09-134-000C-2719	Sequence 2719, App
2	1881	100.0	1983	US-09-071-035-217	Sequence 217, App
3	1881	100.0	1983	US-10-206-576-217	Sequence 217, App
4	1881	89.7	1687	US-09-071-035-219	Sequence 219, App
5	1687	89.7	1687	US-10-206-576-219	Sequence 219, App
6	1014.6	53.9	1995	US-09-107-532A-3598	Sequence 3598, App
7	74.6	4.0	7218	US-08-232-463-14	Sequence 14, App
8	64.6	3.4	6609	US-10-172-502-1	Sequence 1280, App
9	57.2	3.0	1039	US-09-902-540-1880	Sequence 1357, App
10	57	3.0	612	US-08-956-171E-171	Sequence 171, App
11	56.6	3.0	11126	US-08-781-986A-171	Sequence 171, App
12	56.6	3.0	1664976	US-08-916-421B-1	Sequence 1, App
13	56.4	3.0	1664976	US-09-692-570-1	Sequence 22, App
14	54	2.9	1141	US-09-806-708B-22	Sequence 2718, App
15	53.8	2.9	1434	US-09-134-000C-2718	Sequence 50, App
16	53	2.8	3942	US-09-601-198-50	Sequence 2, App
17	53.2	2.8	5361	US-08-973-462-2	Sequence 1, App
18	52.2	2.8	6152	US-08-973-462-1	Sequence 15487, App
19	52.2	2.8	18016	US-09-949-016-15487	Sequence 339, App
20	52	2.8	1898	US-09-769-787-339	Sequence 54, App
21	52	2.8	20986	US-08-961-527-54	Sequence 133019, App
22	51.8	2.8	601	US-09-949-016-133019	

24	51.8	2.8	1891	3	US-08-973-462-3	Sequence 3, App
25	51.6	2.7	614	2	US-09-902-540-1318	Sequence 1318, App
26	51.6	2.7	5181	2	US-08-257-073-10	Sequence 10, App
27	51	2.7	2223	2	US-08-257-073-4	Sequence 4, App
28	51	2.7	640681	3	US-09-790-988-1	Sequence 1, App
29	50.8	2.7	963	3	US-09-134-000C-2132	Sequence 2132, App
30	50.6	2.7	1305	3	US-09-329-234A-6	Sequence 6, App
31	50.4	2.7	1664976	3	US-08-916-421B-1	Sequence 1, App
32	50.4	2.7	1664976	3	US-09-692-570-1	Sequence 1, App
33	50	2.7	1353	3	US-08-956-171E-521	Sequence 521, App
34	50	2.7	1353	3	US-08-781-986A-521	Sequence 521, App
35	49.8	2.6	699	3	US-09-461-697-193	Sequence 193, App
36	49.8	2.6	699	3	US-09-461-697-191	Sequence 191, App
37	49.8	2.6	717	3	US-09-461-697-189	Sequence 189, App
38	49.8	2.6	774	3	US-09-461-697-187	Sequence 187, App
39	49.8	2.6	819	3	US-09-461-697-185	Sequence 185, App
40	49.8	2.6	1669	3	US-09-461-697-184	Sequence 184, App
41	49.8	2.6	1983	3	US-09-107-532A-3061	Sequence 3061, App
42	49.8	2.6	50000	3	US-09-662-254B-25	Sequence 25, App
43	49.6	2.6	5887	4	US-09-747-385-6	Sequence 6, App
44	49.6	2.6	9955	4	US-09-747-385-15	Sequence 15, App
45	49	2.6	1116	3	US-09-248-796A-9660	Sequence 9660, App

ALIGNMENTS

RESULT 1
US-09-134-000C-2719
; Sequence 2719, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2719
; LENGTH: 1980
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-2719

Query Match 100.0%; Score 1881; DB 3; Length 1980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAAGCAATTAAAAAGTTGGTACACCGTTAGTACCTGTATCTAATTTGCCACTT	60
DB	97	ATGAAGCAATTAAAAAGTTGGTACACCGTTAGTACCTGTATCTAATTTGCCACTT	156
QY	61	TTTCAAGTGTATTGGGACCAACCTGATTTGCAAGAAAGGGAGCGGACAG	120
DB	157	TTTCAAGTGTATTGGGACCAACCTGATTTGCAAGAAAGGGAGCGGACAG	216
QY	121	CTCGGATTTCAAAAAGAAATGACGATTTACAGATCCGCTTATTTCAATATGCGGG	180
DB	217	CTCGGATTTCAAAAAGAAATGACGATTTACAGATCCGCTTATTTCAATATGCGGG	276
QY	181	AAAGAAATGAGCGATTTGATTAATATGACGATGGAATGATGATGATTTAT	240
DB	277	AAAGAAATGAGCGATTTGATTAATATGACGATGGAATGATGATGATTTAT	336
QY	241	AAAGTGAAGCAAGATTTTACAGACAGAGCGGACGCAAGGTTGATGACACTAA	300
DB	337	AAAGTGAAGCAAGATTTTACAGACAGAGCGGACGCAAGGTTGATGACACTAA	396
QY	301	CAAGCTGTCAAAAGTTTAATCTCTGGAAGCAACTGTGCTCAAGGAACGATCAAT	360

Db 387 CAAGCTGTCGAAAGTTTAACTCTCGGAAACCTGTTGCTCAAGAAACCGATGCCAAT 456
Qy 361 GGGAAATGCTACTGTTCACTTACCTTAAACAAATGTTAAAGATGCACTGTATACAT 420
Db 457 GGGAAATGCTACTGTTCACTTACCTTAAACAAATGTTAAAGATGCACTGTATACAT 516
Qy 421 AAAGAAACCAAAAGAGGTGTAGTGTCTAGGAATATNGTGTGTGCGCTTCCAGTT 480
Db 517 AAAGAAACCAAAAGAGGTGTAGTGTCTAGGAATATNGTGTGTGCGCTTCCAGTT 576
Qy 481 TACGAATGATCAACAAACAGATGTTCTTAAATATGAAACAGAAATTAAGCGTT 540
Db 577 TACGAATGATCAACAAACAGATGTTCTTAAATATGAAACAGAAATTAAGCGTT 636
Qy 541 GTTCATATTTATCTTAAATATGTTAGTGTAGCCAAATGATGTTTACATGTAAAAAGTA 600
Db 637 GTTCATATTTATCTTAAATATGTTAGTGTAGCCAAATGATGTTTACATGTAAAAAGTA 696
Qy 601 GGAATCTGCTGAATAAGAGATTTAAATGGCGAGAAATTTGTTATTTCTTAAAGCGAAGC 660
Db 697 GGAATCTGCTGAATAAGAGATTTAAATGGCGAGAAATTTGTTATTTCTTAAAGCGAAGC 756
Qy 661 TCACCAAGGACAGTAAATATATCCAGAGATCAAAAGATGATTAATATATCATGCAACG 720
Db 757 TCACCAAGGACAGTAAATATATCCAGAGATCAAAAGATGATTAATATATCATGCAACG 816
Qy 721 GATTAAGAACAGCAAAACGCTTTATTAATCTGGGAAAAGTTATGAAATGGCCAAATATGAT 780
Db 817 GATTAAGAACAGCAAAACGCTTTATTAATCTGGGAAAAGTTATGAAATGGCCAAATATGAT 876
Qy 781 TTCACAGAGCAGAGATGGAACGGGAGAAATTAACGTTAAATAATCTTGAGGTGGTTG 840
Db 877 TTCACAGAGCAGAGATGGAACGGGAGAAATTAACGTTAAATAATCTTGAGGTGGTTG 936
Qy 841 TATATTTTGAAGAGTAAAGCTCCAAATTAATGCAATTAATGAAATCAACAAA 900
Db 937 TATATTTTGAAGAGTAAAGCTCCAAATTAATGCAATTAATGAAATCAACAAA 996
Qy 901 ACACCATTTCAATTTGAGCAAAACATCAAAACCTGTTGAAAACAGTCAAAATGAT 960
Db 997 ACACCATTTCAATTTGAGCAAAACATCAAAACCTGTTGAAAACAGTCAAAATGAT 1056
Qy 961 ACCTCTAAAGTTGATTAACCAACACCAACCTTGAATGTTAAAGATGTCGAATGGCGAA 1020
Db 1057 ACCTCTAAAGTTGATTAACCAACACCAACCTTGAATGTTAAAGATGTCGAATGGCGAA 1116
Qy 1021 AAAATTTAATATCAAAATTTCTGTAATATTTCAATGGGATTTGCACAAAGAGCGAC 1080
Db 1117 AAAATTTAATATCAAAATTTCTGTAATATTTCAATGGGATTTGCACAAAGAGCGAC 1176
Qy 1081 GCTAATTAATACGTCAAAATTTCAATTTGATTAACAGATGACAGCTTAACTTTGAT 1140
Db 1177 GCTAATTAATACGTCAAAATTTCAATTTGATTAACAGATGACAGCTTAACTTTGAT 1236
Qy 1141 AACGTGACTTCTGAGAGATGATGCTTAATGCGTTATATGATGGGATCAGATGTCCT 1200
Db 1237 AACGTGACTTCTGAGAGATGATGCTTAATGCGTTATATGATGGGATCAGATGTCCT 1296
Qy 1201 GAAAATTTACAGATGATGTAACAGCAAAATGCGCTTCACTGTGCGCTTAATCCAGGAT 1260
Db 1297 GAAAATTTACAGATGATGTAACAGCAAAATGCGCTTCACTGTGCGCTTAAATCCAGGAT 1356
Qy 1261 ATTCTTACGCTTAACACAGAGCGGACACTTAAATTTGTTTACTTTATGATTTAAATGAA 1320
Db 1357 ATTCTTACGCTTAACACAGAGCGGACACTTAAATTTGTTTACTTTATGATTTAAATGAA 1416
Qy 1321 AAAGCAGATCTTACGAAAGGCTTTAAATAAGAGCAATGTTGATAACGCTATACCGAC 1380
Db 1417 AAAGCAGATCTTACGAAAGGCTTTAAATAAGAGCAATGTTGATAACGCTATACCGAC 1476
Qy 1381 GACCAAAACCAACCACTGTGTAAGTTGTGACAGGTGGGAAACGTTTCAATTAAGTCAT 1440

Db 1477 GACCAAAACCAACCACTGTGTAAGTTGTGACAGGTGGGAAACGTTTCAATTAAGTCAT 1536
Qy 1441 GGCATGTCACAGCACAACAGCTTGGCGGAGCTTCTTGTGTCCGTGATCAAAAC 1500
Db 1537 GGCATGTCACAGCACAACAGCTTGGCGGAGCTTCTTGTGTCCGTGATCAAAAC 1596
Qy 1501 AGCGACAGCAAAATTTATTTGAAATTCGATGAAACAGAAAGCAGCAACTGGTGAAA 1560
Db 1597 AGCGACAGCAAAATTTATTTGAAATTCGATGAAACAGAAAGCAGCAACTGGTGAAA 1656
Qy 1561 ACAAAGCTGAAGCACTACTTTTAAACAACCGCTGATGATTAATGATATCAAGGG 1620
Db 1657 ACAAAGCTGAAGCACTACTTTTAAACAACCGCTGATGATTAATGATATCAAGGG 1716
Qy 1621 CTTAATATCGGTACTATTTATTAAGAAATCTGAGTCCGATGATTAATGATGTTA 1680
Db 1717 CTTAATATCGGTACTATTTATTAAGAAATCTGAGTCCGATGATTAATGATGTTA 1776
Qy 1681 ACAAATCGGATTTGATTTGTGTCATGTAACATCATATGACCAACAGAAACCTAGTT 1740
Db 1777 ACAAATCGGATTTGATTTGTGTCATGTAACATCATATGACCAACAGAAACCTAGTT 1836
Qy 1741 TCACCAAGAAAAGTACCAACCAACCAAGGTACTTACTTCAACAGGTGGCAAGGA 1800
Db 1837 TCACCAAGAAAAGTACCAACCAACCAAGGTACTTACTTCAACAGGTGGCAAGGA 1896
Qy 1801 ATCTACGTTTCTTGAAGGTGGCGAGCTTGTCTACTTATTTGAGAGTCTACTTGT 1860
Db 1897 ATCTACGTTTCTTGAAGGTGGCGAGCTTGTCTACTTATTTGAGAGTCTACTTGT 1956
Qy 1861 AGACGTAGAAAAGAAAATGCT 1881
Db 1957 AGACGTAGAAAAGAAAATGCT 1977

RESULT 2
US-09-071-035-217
; Sequence 217, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 217:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1983 base pairs
; TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
US-09-071-035-217

Query Match 100.0%; Score 1881; DB 3; Length 1983;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AAGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTTGTACTAATTTTGGCACTT 60
DB 100 AAGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTTGTACTAATTTTGGCACTT 159
QY 61 TTCAAGATGATTAAGGACAACAACCTGATTTGAGAAGAAATGGGAGAGCGACAG 120
DB 160 TTCAAGATGATTAAGGACAACAACCTGATTTGAGAAGAAATGGGAGAGCGACAG 219
QY 121 CTGCGATTCACAAAAGAAAATGACGATTTACAGATCCGCTTATTCAAAATAGCGGG 180
DB 220 CTGCGATTCACAAAAGAAAATGACGATTTACAGATCCGCTTATTCAAAATAGCGGG 279
QY 181 AAGAAGATGAGGAGTTGATTAATATCAAGATCTGGCAGATGTCACGTTTATTTAT 240
DB 280 AAGAAGATGAGGAGTTGATTAATATCAAGATCTGGCAGATGTCACGTTTATTTAT 339
QY 241 AACGTGACGAACGAATTTTACGAGCAACGAGCGGACGAGCGTATGATGACGCTTAA 300
DB 340 AACGTGACGAACGAATTTTACGAGCAACGAGCGGACGAGCGTATGATGACGCTTAA 399
QY 301 CAAGCTGTCGAAGTTTAACTCTGAGGAAACCTGTTGCTCAAGAAACACCGATGCAAT 360
DB 400 CAAGCTGTCGAAGTTTAACTCTGAGGAAACCTGTTGCTCAAGAAACACCGATGCAAT 459
QY 361 GGAATGTCACCTGTCAGTTACTTAAAAAACAATATGTAAGATGCAATGATACAT 420
DB 460 GGAATGTCACCTGTCAGTTACTTAAAAAACAATATGTAAGATGCAATGATACAT 519
QY 421 AAGAAGAACCAAAAGAGGTGATGTCGCTACGAATATGATGATGATGATGATGAT 480
DB 520 AAGAAGAACCAAAAGAGGTGATGTCGCTACGAATATGATGATGATGATGATGAT 579
QY 481 TACGAATGATCAAGCAACAGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 580 TACGAATGATCAAGCAACAGATGATGATGATGATGATGATGATGATGATGATGAT 639
QY 541 GTTCATATTTATCTTAAAAATGTCGATGCAATGATGATGATGATGATGATGATGAT 600
DB 640 GTTCATATTTATCTTAAAAATGTCGATGCAATGATGATGATGATGATGATGATGAT 699
QY 601 GGAATGCTGAATAATGAAGATTAATAGCGCAGAAATTTGTTATTTCTAAAAAGCAAG 660
DB 700 GGAATGCTGAATAATGAAGATTAATAGCGCAGAAATTTGTTATTTCTAAAAAGCAAG 759
QY 661 TCACCAAGCACAAGTAAATATATCCAGAGATCAAGATGATGATGATGATGATGATGAT 720
DB 760 TCACCAAGCACAAGTAAATATATCCAGAGATCAAGATGATGATGATGATGATGATGAT 819
QY 721 GATTAAGAAACAAGAAAACGCTTTATTAATGAGGAAAATGATGATGATGATGATGAT 780
DB 820 GATTAAGAAACAAGAAAACGCTTTATTAATGAGGAAAATGATGATGATGATGATGAT 879
QY 781 TTCAAGAGAGAGATGAGAACGAGGAGAAATTAACATGATTAATTAATTAATTAATTAAT 840
DB 880 TTCAAGAGAGAGATGAGAACGAGGAGAAATTAACATGATTAATTAATTAATTAATTAAT 939
QY 841 TATATTTAGAGAGATTAAGATCAATATATGCAATTAATTAATTAATTAATTAATTAAT 900
DB 940 TATATTTAGAGAGATTAAGATCAATATATGCAATTAATTAATTAATTAATTAATTAAT 999
QY 901 AACACATTTACAAATGAGAGCAACAATCAACCTGTTGAGAAAACAGTCAAAAATAT 960
DB 1000 AACACATTTACAAATGAGAGCAACAATCAACCTGTTGAGAAAACAGTCAAAAATAT 1059
QY 961 ACCTCTAAAGTTGATTAACAACAACCAAGCTTATGATGATGATGATGATGATGATGAT 1020
```

```
DB 1060 ACCTCTAAAGTTGATTAACAACAACCAAGCTTATGATGATGATGATGATGATGATGAT 1119
QY 1021 AAAATTAATTAATCAAAATTTCTGTAATATATTCATTTGGGATGTCAGACAAAGAGCGAC 1080
DB 1120 AAAATTAATTAATCAAAATTTCTGTAATATATTCATTTGGGATGTCAGACAAAGAGCGAC 1179
QY 1081 GCTAATTAATTAATCAAAATTTCTGTAATATATTCATTTGGGATGTCAGACAAAGAGCGAC 1140
DB 1180 GCTAATTAATTAATCAAAATTTCTGTAATATATTCATTTGGGATGTCAGACAAAGAGCGAC 1239
QY 1141 AACGTGACTTCGAGAGATATGCTTATGCTTATATGATGATGATGATGATGATGATGAT 1200
DB 1240 AACGTGACTTCGAGAGATATGCTTATGCTTATATGATGATGATGATGATGATGATGAT 1299
QY 1201 GAAAATTAATCAAGTATGTAACCAACAAATGCTTCACTGTCGCTTATATCAAGCTAT 1260
DB 1300 GAAAATTAATCAAGTATGTAACCAACAAATGCTTCACTGTCGCTTATATCAAGCTAT 1359
QY 1261 ATTCTAGCTTACACAGGCGGCAACCTAAATTCGTTACTTATGATTAATTAATGAT 1320
DB 1360 ATTCTAGCTTACACAGGCGGCAACCTAAATTCGTTACTTATGATTAATTAATGAT 1419
QY 1321 AAGAAGATCTTACGAAGAGCTTTTAAAAATGAGGGAATGTTGATTAACGTCATACCGAC 1380
DB 1420 AAGAAGATCTTACGAAGAGCTTTTAAAAATGAGGGAATGTTGATTAACGTCATACCGAC 1479
QY 1381 GACCAACACCAACCACTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1480 GACCAACACCAACCACTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1539
QY 1441 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1540 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1599
QY 1501 AGCGACACAGCAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1600 AGCGACACAGCAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1659
QY 1561 ACAAAAGCTGAAGCACTTATTAACAACAACGCTGATGATGATGATGATGATGATGAT 1620
DB 1660 ACAAAAGCTGAAGCACTTATTAACAACAACGCTGATGATGATGATGATGATGATGAT 1719
QY 1621 CTTAATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
DB 1720 CTTAATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779
QY 1681 ACAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1780 ACAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1839
QY 1741 TCACCAAGAAAAGTACCAACAACAACAGGATCCTTACCTTCAACAGTGGCAAGGA 1800
DB 1840 TCACCAAGAAAAGTACCAACAACAACAGGATCCTTACCTTCAACAGTGGCAAGGA 1899
QY 1801 ATCTAGCTTCTTATGAGAGGCGCAGCTTCTGATCTTATGAGAGGATGATGATGAT 1860
DB 1900 ATCTAGCTTCTTATGAGAGGCGCAGCTTCTGATCTTATGAGAGGATGATGATGAT 1959
QY 1861 AGACGTAGAAAAGAAATGCT 1881
DB 1960 AGACGTAGAAAAGAAATGCT 1980
```

RESULT 3
US-10-206-576-217
; Sequence 217, Application us/10206576
; Patent No. 6913907
GENERAL INFORMATION:
APPLICANT: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:

ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/206,576
FILING DATE: 29-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB869PDI
INFORMATION FOR SEQ ID NO: 217:
SEQUENCE CHARACTERISTICS:
LENGTH: 1983 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 217:
US-10-206-576-217
Query Match 100.0%; Score 1881; DB 3; Length 1983;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTTTGTAATAATTTGGCAGTT	60
DB	100	ATGAAGCAATTAATAAAGTTGGTACACCGTTAGTACCTTTGTAATAATTTGGCAGTT	159
QY	61	TTCAAGAGTATTTAGGGACAACAAGTTTGGAGAAATTTGGGAGAGCGCAG	120
DB	160	TTCAAGAGTATTTAGGGACAACAAGTTTGGAGAAATTTGGGAGAGCGCAG	219
QY	121	CTGGTATTCACAAAAGAAATGACGATTTACAGATCCGCTTATTCAAAATAGCGG	180
DB	220	CTGGTATTCACAAAAGAAATGACGATTTACAGATCCGCTTATTCAAAATAGCGG	279
QY	181	AAAGAAATGACGAGTTTATTAATATCAAGACCTGGCAGATGTGACGTTTATTTAT	240
DB	280	AAAGAAATGACGAGTTTATTAATATCAAGACCTGGCAGATGTGACGTTTATTTAT	339
QY	241	AACGTGACGAAGATTTTACGAGCAAGAGCGGCGCAAGCCTTATGACGTAATA	300
DB	340	AACGTGACGAAGATTTTACGAGCAAGAGCGGCGCAAGCCTTATGACGTAATA	399
QY	301	CAAGCTGTCAGAAAGTTTAACTCTGGGAAACCTGTTGCTCAAGGAACACCGATGAAT	360
DB	400	CAAGCTGTCAGAAAGTTTAACTCTGGGAAACCTGTTGCTCAAGGAACACCGATGAAT	459
QY	361	GGGAATGTCACTGTTCACTTAATAAACAATATGTAAGATGCAATGATATACATT	420
DB	460	GGGAATGTCACTGTTCACTTAATAAACAATATGTAAGATGCAATGATATACATT	519
QY	421	AAAGAAAGCAAAAGAGGTGTAGTTGCTGCTACGAATATGTTGGTGGCGTTCCAGTT	480
DB	520	AAAGAAAGCAAAAGAGGTGTAGTTGCTGCTACGAATATGTTGGTGGCGTTCCAGTT	579

QY	481	TACGAATGATCAAGCAAAAGATGTTCTTAATAATATGGAACAGAAATTAACGGTT	540
DB	580	TACGAATGATCAAGCAAAAGATGTTCTTAATAATATGGAACAGAAATTAACGGTT	639
QY	541	GTTCAATTTTATCTTAATAAATGTGTAGCCAAATGATGTTTACATGTGAAAAAGTA	600
DB	640	GTTCAATTTTATCTTAATAAATGTGTAGCCAAATGATGTTTACATGTGAAAAAGTA	699
QY	601	GGAACTGCTGAAAAATGAAGATTAATATGGCGGAGAAATTTGTTATTTCTAAAAAGC	660
DB	700	GGAACTGCTGAAAAATGAAGATTAATATGGCGGAGAAATTTGTTATTTCTAAAAAGC	759
QY	661	TCACGAGCAAGTAATAATATATCCAGAGAGTCAAGAATGATTTATATATGCAACG	720
DB	760	TCACGAGCAAGTAATAATATATCCAGAGAGTCAAGAATGATTTATATATGCAACG	819
QY	721	GATTAAGAACAGCAAAACGCTTTATTAATCTGGGAAAGTTATGAATTTGGCAAAATGAT	780
DB	820	GATTAAGAACAGCAAAACGCTTTATTAATCTGGGAAAGTTATGAATTTGGCAAAATGAT	879
QY	781	TTCAAGAGAGCAAGATGGAACGGGAAATTAACAGTTAAAAATCTGAGGTTGCTCG	840
DB	880	TTCAAGAGAGCAAGATGGAACGGGAAATTAACAGTTAAAAATCTGAGGTTGCTCG	939
QY	841	TATATTTTGAAGAGTAATAAGCTCCAAATATATGAGATTAATTAATGAATAATCAAAA	900
DB	940	TATATTTTGAAGAGTAATAAGCTCCAAATATATGAGATTAATTAATGAATAATCAAAA	999
QY	901	ACACCAATTTACATTTGAAGCAACATCAACACTGTTGAAAAACAGTCAAAATGAT	960
DB	1000	ACACCAATTTACATTTGAAGCAACATCAACACTGTTGAAAAACAGTCAAAATGAT	1059
QY	961	ACCTTAAGTTGATTAATAACAACCAAGCTTAGATGGTAAAGATGTGGCAATGGCGAA	1020
DB	1060	ACCTTAAGTTGATTAATAACAACCAAGCTTAGATGGTAAAGATGTGGCAATGGCGAA	1119
QY	1021	AAAAATTAATATCAAAATTTCTGTAAATATTCATTTGGGAAATTCAGAACAGGCAAC	1080
DB	1120	AAAAATTAATATCAAAATTTCTGTAAATATTCATTTGGGAAATTCAGAACAGGCAAC	1179
QY	1081	GCTAATTAATATCTCAATTTCAATTTAGTTGATTAACATGATGACGCTTAATCTTTGAT	1140
DB	1180	GCTAATTAATATCTCAATTTCAATTTAGTTGATTAACATGATGACGCTTAATCTTTGAT	1239
QY	1141	AACGTGACTTCGAGAGATGATGCTTATGCTATATGATGAGGATACAGTATGCTCT	1200
DB	1240	AACGTGACTTCGAGAGATGATGCTTATGCTATATGATGAGGATACAGTATGCTCT	1299
QY	1201	GAAAAATTAATCAAGTATGAAACAAACAAATGGCTTCACTGTCGCGTTAATCCAGCTAT	1260
DB	1300	GAAAAATTAATCAAGTATGAAACAAACAAATGGCTTCACTGTCGCGTTAATCCAGCTAT	1359
QY	1261	ATTCTAGGCTTAACACGAGCGGCACTTAATAATTCGTTTACTTTATGATTAATGAA	1320
DB	1360	ATTCTAGGCTTAACACGAGCGGCACTTAATAATTCGTTTACTTTATGATTAATGAA	1419
QY	1321	AAAGCAGATCTTACGAAAGGCTTTTAAAAATGAGCGAAATGTTGATTAACGTCATACCGAC	1380
DB	1420	AAAGCAGATCTTACGAAAGGCTTTTAAAAATGAGCGAAATGTTGATTAACGTCATACCGAC	1479
QY	1381	GACCAAAACCAACCAACTGTGTAAGTTGACAGGTGGGAAAGCTTTCAATTAAGTCGAT	1440
DB	1480	GACCAAAACCAACCAACTGTGTAAGTTGACAGGTGGGAAAGCTTTCAATTAAGTCGAT	1539
QY	1441	GGCGATGTGACAGCGACACAGCCTTTGGGGAGCTTCTTTGTGTCGCTGATCAAAAC	1500
DB	1540	GGCGATGTGACAGCGACACAGCCTTTGGGGAGCTTCTTTGTGTCGCTGATCAAAAC	1599
QY	1501	ACGCAACAGCAAAATTAATTTGAAAAATCGATGAACAAACGAGCAACTTGGGTGAAA	1560
DB	1600	ACGCAACAGCAAAATTAATTTGAAAAATCGATGAACAAACGAGCAACTTGGGTGAAA	1659
QY	1561	ACAAAAAGCTGAAGCAACTTATCTTTTACAAACAGCGCTGATGATGATTGATATCACAGG	1620

Accession	Sequence	Position
D8	1660 ACAAAAGCTGAGCAACTCTTTTACAAACAACGGCTGATGATTAGTTGATATCAACGGG	1719
QY	1621 CTTAAATACGGTACCTATTTATTTAGAAAGAACTGAGCTCCGTGATTTATGTCTTTTA	1680
D8	1720 CTTAAATACGGTACTCTATTTATTTAGAAAGAACTGAGCTCCTGATGATTAATGTCTTTTA	1779
QY	1681 ACAAAATCGGATTTGAATTTTGTGTGCAATGAAACAATCATATGTGCACAACAGAAAACCTTAGTT	1740
D8	1780 ACAATTCGATTTGAATTTTGTGTGCAATGAAACAATCATATGTGCACAACAGAAAACCTTAGTT	1839
QY	1741 TCACACAGAAAAAGTACCAACAAACACAACAGTACTTACCTTCACACAGGTGGCAAAAGGA	1800
D8	1840 TCACACAGAAAAAGTACCAACAAACACAACAGTACTTACCTTCACACAGGTGGCAAAAGGA	1899
QY	1801 ATCTACGTTTATCTTAGGAAGTGGCGCACTCTTGTACTTATTTGACAGAGTCTACTTTTGT	1860
D8	1900 ATCTACGTTTACTTAGGAAGTGGCGCACTCTTGTACTTATTTGACAGAGTCTACTTTTGT	1959
QY	1861 AGACGTAGAAAAAGAAAATGCT 1881	
D8	1960 AGACGTAGAAAAAGAAAATGCT 1980	

RESULT 4

```

US-09-071-035-219
; Sequence 219, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1687 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-071-035-219

```

	Query Match	Similarity	89.7%	Score 1687	DB 3	Length 1687
	Best Local	Similarity	100.0%	Pred. No. 0		
	Matches 1687	Conservative	0	Mismatches 0	Indels 0	Gaps 0
Qy	96	AGAAAGAAATATGCGAGACGCGACAGCTCTGTGATTACAAAAGAAATATGACGATTTACC	155			
Qb	1	AGAAAGAAATATGCGAGACGCGACAGCTCTGTGATTACAAAAGAAATATGACGATTTACC	60			

QY	156	AGATCCGCTTTTTCAAAAATACCGGGAAAGAAATGACGAGCTTGATAAATTCAGAGCT	215
Db	61	AGATCCGCTTTTTCAAAAATACCGGGAAAGAAATGACGAGCTTGATAAATTCAGAGCT	120
QY	216	GGCAGATGTGACGTTTAGTATTTATACGTGACGACAGATTTTATCGACGACGCGC	275
Db	121	GGCAGATGTGACGTTTAGTATTTATACGTGACGACAGATTTTATCGACGACGCGC	180
QY	276	AGGCGCAGCGTTGATGACGCTAAACAAGCTGTCCAAAGTTTAACTCTGGGAAACTGT	335
Db	181	AGGCGCAGCGTTGATGACGCTAAACAAGCTGTCCAAAGTTTAACTCTGGGAAACTGT	240
QY	336	TGCTCAAGAAACCAACCGATGCAAAATGGGAATGTCACTGTCAAGTACTTAAAAACAAA	395
Db	241	TGCTCAAGAAACCAACCGATGCAAAATGGGAATGTCACTGTCAAGTACTTAAAAACAAA	300
QY	396	TGCTAAAGATGCAAGTGTATACCATTTAAAGAAACCAAAAGGGGTGATGTGCTGCTAC	455
Db	301	TGCTAAAGATGCAAGTGTATACCATTTAAAGAAACCAAAAGGGGTGATGTGCTGCTAC	360
QY	456	GAATATGTTGTGTGGCGTTCCAGTTTACGAATGATCAAGCAACAGTTGTTCTTATA	515
Db	361	GAATATGTTGTGTGGCGTTCCAGTTTACGAATGATCAAGCAACAGTTGTTCTTATA	420
QY	516	ATTATGACACAAAGAAATTAAGCGTGTGTCATATTATTCCTAAAAATGTGTAGCCATAGA	575
Db	421	ATTATGACACAAAGAAATTAAGCGTGTGTCATATTATTCCTAAAAATGTGTAGCCATAGA	480
QY	576	TGTATGTTTACATGTGAAAAAATGATGGAACTGTGAAAAATGAGAGTTAAATGGCGCAGA	635
Db	481	TGTATGTTTACATGTGAAAAAATGATGGAACTGTGAAAAATGAGAGTTAAATGGCGCAGA	540
QY	636	ATTGTGTTATTTCTAAAAAGCGAAGGCTCACAGGCAAGTAAATATATCCAAAGAGTCAA	695
Db	541	ATTGTGTTATTTCTAAAAAGCGAAGGCTCACAGGCAAGTAAATATATCCAAAGAGTCAA	600
QY	696	AGATGATTTATATCATGGAACAACGATTAAGAAACAAGCAAAACGCTTATTACTGGGAA	755
Db	601	AGATGATTTATATCATGGAACAACGATTAAGAAACAAGCAAAACGCTTATTACTGGGAA	660
QY	756	AAGTTATGAAATTTGGCGAAAATGATTTCAACGAAGCAGAGAAATGGAAACGGAGAAATTAC	815
Db	661	AAGTTATGAAATTTGGCGAAAATGATTTCAACGAAGCAGAGAAATGGAAACGGAGAAATTAC	720
QY	816	AGTTAAAAAATTTGAGGTTGTTGTTATTTTATGAAAGATAAAGTCCAAATTAATGC	875
Db	721	AGTTAAAAAATTTGAGGTTGTTGTTATTTTATGAAAGATAAAGTCCAAATTAATGC	780
QY	876	AGAAATTATTTGAAATCAAAACAACCAATTTACATTTGAGCAACCAATCAACACCC	935
Db	781	AGAAATTATTTGAAATCAAAACAACCAATTTACATTTGAGCAACCAATCAACACCC	840
QY	936	TGTTGAAAAAACAATCAAAAAATGATACCTCTAAAGTTGATTAACCAACCAAGGTTTGA	995
Db	841	TGTTGAAAAAACAATCAAAAAATGATACCTCTAAAGTTGATTAACCAACCAAGGTTTGA	900
QY	996	TGCTAAAGATGTGCAATTGGCGAAAAAATTTAAATATCAAAATTTCTGTAAATATTCACAT	1055
Db	901	TGCTAAAGATGTGCAATTGGCGAAAAAATTTAAATATCAAAATTTCTGTAAATATTCACAT	960
QY	1056	GGGATTTGACAGCAAAAGAGCGACGCTAATAAATAGTCAAAATTCATTTAGTTGATTA	1115
Db	961	GGGATTTGACAGCAAAAGAGCGACGCTAATAAATAGTCAAAATTCATTTAGTTGATTA	1020
QY	1116	ACATGATGACGCTTAACTTTTGAATACGTGACTTGTGAGAGATGTCTTATGCGTTATA	1175
Db	1021	ACATGATGACGCTTAACTTTTGAATACGTGACTTGTGAGAGATGTCTTATGCGTTATA	1080
QY	1176	TGATGGGATATCAGTGATGTGCTCTGAAAAATTAATCAAGTGAATGACAAATGGCTT	1235
Db	1081	TGATGGGATATCAGTGATGTGCTCTGAAAAATTAATCAAGTGAATGACAAATGGCTT	1140

QY	1236	CACCTGCGCGCTTAATCCAGCGGTATATCTTCAAGCTAAACACAGCGGCGACACTTAAATT	1235
Db	1141	CACCTGCGCGCTTAATCCAGCGGTATATCTTCAAGCTAAACACAGCGGCGACACTTAAATT	1200
QY	1296	CGTTTACTTATGTCATTTAAATGAAAAAGAGACTCTTACGAAAGGCTTTAAAAATGAGGC	1355
Db	1201	CGTTTACTTATGTCATTTAAATGAAAAAGAGACTCTTACGAAAGGCTTTAAAAATGAGGC	1260
QY	1356	GAAATGTTGATTAACGGGTATACCGACGACCAACCAACCAACTGTTGAAGTTGTGACAGG	1415
Db	1261	GAAATGTTGATTAACGGGTATACCGACGACCAACCAACCAACTGTTGAAGTTGTGACAGG	1320
QY	1416	TGGGAAACGTTTCATTAAAGTCATGAGCGCATGTGACAGCGCACAGAGCCTTGGCGGAGC	1475
Db	1321	TGGGAAACGTTTCATTAAAGTCATGAGCGCATGTGACAGCGCACAGAGCCTTGGCGGAGC	1380
QY	1476	TTCCCTTGTGTCGCGTGATCAAAAACAGCGCACAGCAAAATTTTGAAGATCGATGAAC	1535
Db	1381	TTCCCTTGTGTCGCGTGATCAAAAACAGCGCACAGCAAAATTTTGAAGATCGATGAAC	1440
QY	1536	AACGAAAGCAGCAACTGGGTGAAAAAACAAGCTGAAGCACTTCTTACAAACGCGC	1595
Db	1441	AACGAAAGCAGCAACTGGGTGAAAAAACAAGCTGAAGCACTTCTTACAAACGCGC	1500
QY	1596	TGATGATTAAGTTGATATCAACAGGCGCTTAAATACGTAACCTATTATTAGAAAGAACTGT	1655
Db	1501	TGATGATTAAGTTGATATCAACAGGCGCTTAAATACGTAACCTATTATTAGAAAGAACTGT	1560
QY	1656	AGCTCCTGATGATTAATGTCTTGTAAACAATCGGATTTGAATTTGGTCAATGAACATC	1715
Db	1561	AGCTCCTGATGATTAATGTCTTGTAAACAATCGGATTTGAATTTGGTCAATGAACATC	1620
QY	1716	ATATGACACAACGAAACCTAGTTTCACGAGAAAGTACCAACCAACCAAGGATC	1775
Db	1621	ATATGACACAACGAAACCTAGTTTCACGAGAAAGTACCAACCAACCAAGGATC	1680
QY	1776	CTTACTT 1782	
Db	1681	CTTACTT 1687	
RESULT 5			
US-10-206-576-219			
; Sequence 219, Application US/10206576			
; Patent No. 6913907			
; GENERAL INFORMATION:			
; APPLICANT: Choi et al.			
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides			
; NUMBER OF SEQUENCES: 497			
; CORRESPONDENCE ADDRESSES:			
; ADDRESSER: Human Genome Sciences, Inc.			
; STREET: 9410 Key West Avenue			
; CITY: Rockville			
; STATE: Maryland			
; COUNTRY: USA			
; ZIP: 20850			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: CD-R			
; COMPUTER: Dell Latitude			
; OPERATING SYSTEM: Windows 98			
; SOFTWARE: ASCII Text			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/10/206,576			
; FILING DATE: 29-Jul-2002			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 09/071,035			
; FILING DATE: 1998-05-04			
; APPLICATION NUMBER: US 60/046,655			
; FILING DATE: 1997-05-16			
; APPLICATION NUMBER: US 60/044,031			
; FILING DATE: 1997-05-06			
; APPLICATION NUMBER: US 60/066,009			

;	FILED DATE: 1997-11-14
;	ATTORNEY/AGENT INFORMATION:
;	NAME: Hyman, Mark J.
;	REGISTRATION NUMBER: 46,789
;	REFERENCE/DOCKET NUMBER: PB369P1D1
;	INFORMATION FOR SEQ ID NO: 219:
;	SEQUENCE CHARACTERISTICS:
;	LENGTH: 1687 base pairs
;	TYPE: nucleic acid
;	STRANDEDNESS: double
;	TOPOLOGY: linear
;	SEQUENCE DESCRIPTION: SEQ ID NO: 219:
US-10-206-576-219	
Query Match	89.7%; Score 1687; DB 3; Length 1687;
Best Local Similarity	100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0
Matches 1687, Conservative	0;
Db	96 AGAAGAAATGCGGAGAGCGCAGCTCGTATTCACAAAAAGAAATGACGGAATTAC 155
Qy	1 AGAAGAAATGCGGAGAGCGCAGCTCGTATTCACAAAAAGAAATGACGGAATTAC 60
Db	156 AGATCCGCTTATTCAAATAGCGGAAAGAAATGACGAGTTGTAATATCAAGACT 215
Qy	61 AGATCCGCTTATTCAAATAGCGGAAAGAAATGACGAGTTGTAATATCAAGACT 120
Db	216 GGCAGATGTGAGTTAGTATTTATTAACGTGACGACGAATTTTACGAGCAACGAGCGC 275
Qy	121 GGCAGATGTGAGTTAGTATTTATTAACGTGACGACGAATTTTACGAGCAACGAGCGC 180
Db	276 AGGCGCAAGCGTTGATGACGCTAAACANAGCTGTCCAAAGTTTAACTCTCGGAAACTGT 335
Qy	181 AGGCGCAAGCGTTGATGACGCTAAACANAGCTGTCCAAAGTTTAACTCTCGGAAACTGT 240
Db	336 TGCTCAAGSAGACCCGATGCAAAATGGGAATGTCACTGTTCAGTTACCTTAACAAAAACAAA 395
Qy	241 TGCTCAAGSAGACCCGATGCAAAATGGGAATGTCACTGTTCAGTTACCTTAACAAAAACAAA 300
Db	396 TGGTAAAGATCAGTGTATACCATTTAAAGAAACCAAAAGGCGTGTAGTTCGTCTAC 455
Qy	301 TGGTAAAGATCAGTGTATACCATTTAAAGAAACCAAAAGGCGTGTAGTTCGTCTAC 360
Db	456 GAATATGTTGGTGGCGTTCCCAAGTTTACGAATGTATCAAGCAAAACAGAGTGTCTTATTA 515
Qy	361 GAATATGTTGGTGGCGTTCCCAAGTTTACGAATGTATCAAGCAAAACAGAGTGTCTTATTA 420
Db	516 ATATGCAACAGAAATGACGAGTGTTCATATTTATCTTAAAAATGTGTACCAATGA 575
Qy	421 ATATGCAACAGAAATGACGAGTGTTCATATTTATCTTAAAAATGTGTACCAATGA 480
Db	576 TGTATGTTTACATGTGAAAAAATGTGAACTGTCTGAAAAATGAGGATTTAAATGGCGCAGA 635
Qy	481 TGTATGTTTACATGTGAAAAAATGTGAACTGTCTGAAAAATGAGGATTTAAATGGCGCAGA 540
Db	636 ATTTGTTATTTCTTAAAGCGAAGCGTCAACGAGCAAGTAAATATATCAAGAGTCAA 695
Qy	541 ATTTGTTATTTCTTAAAGCGAAGCGTCAACGAGCAAGTAAATATATCAAGAGTCAA 600
Db	696 AGATGATTTATATCATGAGCAACGCGATTAAGAAACCAAGCAAAACGCTTTATTACTGGGAA 755
Qy	601 AGATGATTTATATCATGAGCAACGCGATTAAGAAACCAAGCAAAACGCTTTATTACTGGGAA 660
Db	756 AAGTTATGAATTTGGCGGAAATGATTTTCAAGAACACAGAAATGGAACGGGGAATTTAAC 815
Qy	661 AAGTTATGAATTTGGCGGAAATGATTTTCAAGAACACAGAAATGGAACGGGGAATTTAAC 720
Db	816 AGTTAAAAATCTTGAAGTTGGTTCGTATATTTTAAAGAAAGTAAAGCTCCAAATATATGC 875
Qy	721 AGTTAAAAATCTTGAAGTTGGTTCGTATATTTTAAAGAAAGTAAAGCTCCAAATATATGC 780
Db	876 AGAATTAATGAATAATCAAACAAAACCATTTTCAATTTGAAGCAACATCAACACAC 935
Qy	781 AGAATTAATGAATAATCAAACAAAACCATTTTCAATTTGAAGCAACATCAACACAC 840

```
QY 936 TGTGAAAAAAGTCAAAAAATGATACCTCTAAAGTTGATAAACAAACCAAGCTTAGA 995
DB 841 TGTGAAAAAAGTCAAAAAATGATACCTCTAAAGTTGATAAACAAACCAAGCTTAGA 900
QY 996 TGTGAAAGATGTGGCAATTTGGCAAAAAATTAATATCAAAATTTCTGTAAATATTCATT 1055
DB 901 TGTGAAAGATGTGGCAATTTGGCAAAAAATTAATATCAAAATTTCTGTAAATATTCATT 960
QY 1056 GGGGATTTGACAGCAAAAGAGCGCAAGCTTAATTAATACGTCAAAATTCATTTAGTTATTA 1115
DB 961 GGGGATTTGACAGCAAAAGAGCGCAAGCTTAATTAATACGTCAAAATTCATTTAGTTATTA 1020
QY 1116 AACTGATGACAGCTTAACTTTTGTATTAACGTAAGTCTTGAGAGTATGCTTAATCCGTTATA 1175
DB 1021 AACTGATGACAGCTTAACTTTTGTATTAACGTAAGTCTTGAGAGTATGCTTAATCCGTTATA 1080
QY 1176 TGTATGGGATATACAGTATTTGCTCTGAAATTAATCAAGTATGTAACAAGCAAAATGGCTT 1235
DB 1081 TGTATGGGATATACAGTATTTGCTCTGAAATTAATCAAGTATGTAACAAGCAAAATGGCTT 1140
QY 1236 CACTGTGCGCGTTAATCCAGCGTATATTCCTACGCTTAACACAGCGCGCACTATAAATT 1295
DB 1141 CACTGTGCGCGTTAATCCAGCGTATATTCCTACGCTTAACACAGCGCGCACTATAAATT 1200
QY 1296 CGTTTAACTTATGCAATTTAAATGAAGAAAGACAGATCCTACGAAAGGCTTTAAAAATGAGGC 1355
DB 1201 CGTTTAACTTATGCAATTTAAATGAAGAAAGACAGATCCTACGAAAGGCTTTAAAAATGAGGC 1260
QY 1356 GAATGTGTAACAGGTATACAGCAAGCAACCAACCACTGTGTAAGTTGTGACAG 1415
DB 1261 GAATGTGTAACAGGTATACAGCAAGCAACCAACCACTGTGTAAGTTGTGACAG 1320
QY 1416 TGGGAAACGTTTCATTTAAATGATGCGATGTGACAGCGCAACAAGCTTTGGCGGAGC 1475
DB 1321 TGGGAAACGTTTCATTTAAATGATGCGATGTGACAGCGCAACAAGCTTTGGCGGAGC 1380
QY 1476 TTCCTTTGCGTCCGATCAAAAACAGCAACAGCAAAATTTTGAATAATGATGAAC 1535
DB 1381 TTCCTTTGCGTCCGATCAAAAACAGCAACAGCAAAATTTTGAATAATGATGAAC 1440
QY 1536 AACGAAGACAGCACTTGGGTGTAACAAAGCTGAAGCACTACTTTTACAAACAGCGC 1595
DB 1441 AACGAAGACAGCACTTGGGTGTAACAAAGCTGAAGCACTACTTTTACAAACAGCGC 1500
QY 1596 TGATGATTTAGTTGATATATCAAGGCTTAATACGTAATTTTGAAGAAATCTGT 1655
DB 1501 TGATGATTTAGTTGATATATCAAGGCTTAATACGTAATTTTGAAGAAATCTGT 1560
QY 1656 AGCTCCTGATGATTAATGCTTTTAAACAATGGGATTTGCGTCAATGAACAATC 1715
DB 1561 AGCTCCTGATGATTAATGCTTTTAAACAATGGGATTTGCGTCAATGAACAATC 1620
QY 1716 ATATGGCAACAAGAAACCTAGTTTACACAGAAAAAGTACCAACAAACAAAGGTAC 1775
DB 1621 ATATGGCAACAAGAAACCTAGTTTACACAGAAAAAGTACCAACAAACAAAGGTAC 1680
QY 1776 CTTACTT 1782
DB 1681 CTTACTT 1687
```

RESULT 6

```
US-09-107-532A-3598
; Sequence 3598, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
```

```
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: JULY 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3598:
SEQUENCE CHARACTERISTICS:
LENGTH: 1995 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (8) LOCATION 1..1995
SEQUENCE DESCRIPTION: SEQ ID NO: 3598:
US-09-107-532A-3598

Query Match 53.9%; Score 1014.6; DB 3; Length 1995;
Best Local Similarity 71.9%; Pred. No. 5.9e-238;
Matches 1342; Conservative 0; Mismatches 519; Indels 6; Gaps 1;

QY 1 ATGAAGCAATTAATAAAGTTGGTATACACCGTTAGTACCTTGTTACTTAATTTGGCCACTT 60
DB 118 ATGAATAATCATATAAATAAATTAACGTTAGTTAGAGTCTTTCTTCTTAATTTTACCATTA 177
QY 61 TTCAAGTGTATTAGGACAACTGCAATTTGCAAGAAATAATGGGAGAGCGCAG 120
DB 178 CTCACAAACAGCTTGGCGCAAAAAAGTGTGACAGAGACAG-----CAGCTCAA 231
QY 121 CTCGATTTCAAAAAGAAATGAACGATTTTACAGATCCGCTTATTCAAAATAGCGGG 180
DB 232 GTCACTCTTCAATAAAGAAATGACTGATTTTACCGGATCTTTTAATCCAAAACAGCGGG 291
QY 181 AAGAATAAGCGAGTTGATTAATATCAAGGACTGGCAGATGTGACGTTTATATTAT 240
DB 292 AAGAATAAGCGAGTTGATTAATATCAAGGACTGGCAGATGTGACGTTTATATTAT 351
QY 241 AAGTGACAGCAATTTTACAGCAAGAGCGGCAAGCGTATGATGACGTTAA 300
DB 352 AAGTGACAGCAATTTTATGCGCAAGAGGTAAGAGCGTCCGTGATGACAGAAA 411
QY 301 CAGCTGTCAAAAGTTTAACTCTGGGAAACCTGTGCTCAAGGAACACCGATCAAT 360
DB 412 CAGCAGTCAAGCTTTTGACTCTGTGATACCAAGTTGCTTCAAGAAACAGATGCTAT 471
QY 361 GGAATGTCACTGTTCACTTACCTTAATAAACAATGTGTAAGATGACAGTATACAT 420
DB 472 GGAATGTCACTTATCTTTACTTAATAAACAATGTGGAAGATGACAGTATACAGATC 531
```

421 AAAAGAAACCAAAAGGGGTGTAATTGCTAGCAATATGTTGGTGGCTTCCAGTT 480
DB AAAAGAAACCAAAAGGGGTGTAATTGCTAGCAATATGTTGGTGGCTTCCAGTT 591
481 TAAGAAATGATCAAGCAACAGATGGTTCCTATTAATATGAAACAGAAATTAAGCGTT 540
DB TATGATGATCAACAGCAATGGCTCTTATTAATAGGGACAGAAATCTAGTACT 651
541 GTTCAATTATTCCTAAATAATGGTATGCAATGATGTAATGTTTACATGTGAAAAAGTA 600
DB ATCCATCTCTACCTTAAAAATACAGTGGTAAATGAGGAACGTGAAAGTTCAAAAAATC 711
601 GGAACGTGAAAAATGAAAGATTTAAATGGCGCAATTTGTTATTTCTAAAAAGCGAGGC 660
DB GGTATGCGCAAAACAGAACACTAAATGAGCAAGAAATTAATTAATTTCTAAAGAAAGGA 771
661 TCACCAAGGACAGTAATAATATCCAAAGAGTCAAGATGATATATATCATGACAAACG 720
DB AACCAAGCGTCAAAAAATACATCCAAAGTCAAGATGATATGATGACCTTGGACAACT 831
721 GATTAAGAAACCAAGCAACGCTTATTTACTGGAAAAATGTAATTTGGCAAAATGAT 780
DB GATCAAAACCAAGCCCAAACTTCTACTGTCATTTCTTATGACATCGCAACAAATGAC 891
781 TTCAAGAAAGCAAGATGGAACGGGAAATTAACAGTTAAAAATCTTGAGGTGGTTCG 840
DB TTTGCGAGGACATCTATTTAAAAAGCGCAGTTGATGTTATCATTTAAGATTGAAAAA 951
841 TATATTTTGAAGAAAGTAAAGCTCCAAATTAATGCAAAATTAATGAAAAATCAACAAA 900
DB TATATTTTGAAGAAAGTAAAGCTCCCTGTAATGCGAAATGATGAAAAAGCAACAAATC 1011
901 AACCAATTACAAATTGAACCAACCAATCAACCTGTTGAAAAAATCACTCAAAAAATGAT 960
DB AGCCTTTTGAAGATCTGGCAAAATAGCCCAACACCGTGAAGAAAGCCATCAAAAAATGAT 1071
1012 AGCCTTTTGAAGATCTGGCAAAATAGCCCAACACCGTGAAGAAAGCCATCAAAAAATGAT 1071
961 ACCCTAAAGTTGATTAACCAACCAACCAATCAACCTGTTGAAAAATGATGCGCA 1020
DB AGCTTAAGTTGATTAACCAACCAACCAATCAACCTGTTGAAAAATGATGCGCA 1131
1072 AGCTTAAGTTGATTAACCAACCAACCAATCAACCTGTTGAAAAATGATGCGCA 1131
DB AAAATTAATCAAAATTTCTGTAATATTCATTTGGGATTCGACACAAAGAGCGAC 1080
1021 AAAATTAATCAAAATTTCTGTAATATTCATTTGGGATTCGACACAAAGAGCGAC 1080
DB AAAATTAATCAAAATTTCTGTAATATTCATTTGGGATTCGACACAAAGAGCGAC 1191
1132 AAAATTAATCAAAATTTCTGTAATATTCATTTGGGATTCGACACAAAGAGCGAC 1191
1081 GCTAATAATTAAGTCAAAATTTCAATTTAGTTAAACATGATGCGCCTTAATCTTGTAT 1140
DB GCTAATAATTAAGTCAAAATTTCAATTTAGTTAAACATGATGCGCCTTAATCTTGTAT 1251
1192 CAAACCAAGTCAACCAATTTCAATTTAGTTAAACATGATGCGCCTTAATCTTGTAT 1251
DB AACGTGACTTCTGAGAGATGATGCTTATGCTTATATGATGAGGATCACTGATGCTCT 1200
1141 AACGTGACTTCTGAGAGATGATGCTTATGCTTATATGATGAGGATCACTGATGCTCT 1200
DB AATGATTTCTGAGAGATGATGCTTATGCTTATATGATGAGGATCACTGATGCTCT 1311
1252 AATGATTTCTGAGAGATGATGCTTATGCTTATATGATGAGGATCACTGATGCTCT 1311
DB GAAAAATTAATCAAGTCAACCAAGTCACTGCTGCTGCTGCTTAAATCCAGCTAT 1260
1201 GAAAAATTAATCAAGTCAACCAAGTCACTGCTGCTGCTGCTTAAATCCAGCTAT 1260
DB GAAAAATTAATCAAGTCAACCAAGTCACTGCTGCTGCTGCTTAAATCCAGCTAT 1371
1312 GAAAAATTAATCAAGTCAACCAAGTCACTGCTGCTGCTGCTTAAATCCAGCTAT 1371
DB ATTCTAGGCTAAACCAAGGCGGACACTAAATTTGTTATCTTATGATTTAAATGAA 1320
1261 ATTCTAGGCTAAACCAAGGCGGACACTAAATTTGTTATCTTATGATTTAAATGAA 1320
DB ATTCTAGGCTAAACCAAGGCGGACACTAAATTTGTTATCTTATGATTTAAATGAA 1431
1372 ATTCTAGGCTAAACCAAGGCGGACACTAAATTTGTTATCTTATGATTTAAATGAA 1431
DB AAAAGAGATCTTACGAAAGCTTTTAAAAATGAGCGCAATGTTAAACGCTATCCGAC 1380
1321 AAAAGAGATCTTACGAAAGCTTTTAAAAATGAGCGCAATGTTAAACGCTATCCGAC 1380
DB AAAAGAGATCTTACGAAAGCTTTTAAAAATGAGCGCAATGTTAAACGCTATCCGAC 1491
1432 AAAAGAGATCTTACGAAAGCTTTTAAAAATGAGCGCAATGTTAAACGCTATCCGAC 1491
DB GACCAAAACCAACCAATCTTGAAGTTGTAAGTGGGAAACGTTTCACTTAAAGTCAAT 1440
1381 GACCAAAACCAACCAATCTTGAAGTTGTAAGTGGGAAACGTTTCACTTAAAGTCAAT 1440
DB GATCAAAACCAACCGTCACTGATGCTTACTGAGGAGCAAAAGATTTGTTAAAGTCAAT 1551
1492 GATCAAAACCAACCGTCACTGATGCTTACTGAGGAGCAAAAGATTTGTTAAAGTCAAT 1551
DB GCGGATGTAAGCAAGCAAGGCTTGGGAGGAGCTTCCCTTGTGCTGCTATCAAAAC 1500
1441 GCGGATGTAAGCAAGCAAGGCTTGGGAGGAGCTTCCCTTGTGCTGCTATCAAAAC 1500
DB GGTGAGCTTACATCAAGCAAAACCTTGTGAGGAGAAATTCGTCGTCGTCGTCAGAT 1611
1552 GGTGAGCTTACATCAAGCAAAACCTTGTGAGGAGAAATTCGTCGTCGTCGTCAGAT 1611

QY 1501 AGCGACACAGCAAAATTAATTTGAAAAATCGATGAAAAACGAAGCAGCACTTGGTGAAA 1560
DB AGTGAACACAGCAAAATTAATTTGAAAAATCGATGAAAAACGAAGCAGCACTTGGTGAAA 1671
QY 1561 ACMAAAGCTGAAGCACTTCTTTTCAACCAACGCGCTGATGATTTAGTTGATATCAAGGG 1620
DB GCGAAAGATCAGCAACGGTTTTTTTCAACCAACGCGCTGATGATTTAGTTGATATCAAGGG 1731
QY 1621 CTTAATTAACGTAACCTATTAATTTAGAAAGAACTGATGCTCCGATGATTTAGTCTTTGTT 1680
DB CTTAATTAACGTAACCTATTAATTTAGAAAGAACTGATGCTCCGATGATTTAGTCTTTGTT 1791
1732 CTTAATTAACGTAACCTATTAATTTAGAAAGAACTGATGCTCCGATGATTTAGTCTTTGTT 1791
QY 1681 ACMAAAGCTGAAGCACTTCTTTTCAACCAACGCGCTGATGATTTAGTTGATATCAAGGG 1740
DB ACMAAAGCTGAAGCACTTCTTTTCAACCAACGCGCTGATGATTTAGTTGATATCAAGGG 1851
1792 ACMAAAGCTGAAGCACTTCTTTTCAACCAACGCGCTGATGATTTAGTTGATATCAAGGG 1851
QY 1741 TCACCAAGAAAGTACCAACCAACCAACCAAGGATCTTACTTCAACAGGTGGCAAGGA 1800
DB TCCTCTGAAAAAATACCAATTAACCAACCAAGGATCTTACTTCAACAGGTGGCAAGGA 1911
1852 TCCTCTGAAAAAATACCAATTAACCAACCAAGGATCTTACTTCAACAGGTGGCAAGGA 1911
QY 1801 ATCTAGCTTACTTGAAGAGTGGCGGACGCTTCTACTTATTTGAGGAGTCTTACTTGTCT 1860
DB ATCTAGCTTACTTGAAGAGTGGCGGACGCTTCTACTTATTTGAGGAGTCTTACTTGTCT 1971
1912 ATCTAGCTTACTTGAAGAGTGGCGGACGCTTCTACTTATTTGAGGAGTCTTACTTGTCT 1971
QY 1861 AGACGTA 1867
DB 1972 AGACGTA 1978

RESULT 7
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, P.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0239
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT29pt-F18
US-08-232-463-14

Query Match 4.0%; Score 74.6; DB 2; Length 7218;
Best Local Similarity 10.2%; Pred. No. 3.9e-08;

Matches 47; Conservative 221; Mismatches 185; Indels 0; Gaps 0;

```
OY 371 CTGTTCACTTACCTAAAAAACAATGTGAAGATCATGTATACCAATTAAGAAGAAC 430
DB 1519 CTAATTCAGTTTCAAAAAAGCGCATGTAGCATCATGTATATCTATCATGCAAGTAG 1460
OY 431 CAAAAGAGGGTGTAGTGTCTGACGAATATGTGTGCGCTCCAGTTTACGAATAGA 490
DB 1459 TTAAGAGATAGAAAGATTGTGACRCRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1400
OY 491 TCAAGCAACAGATGCTTCTATTAATATGAAACAGAGAATTAAGCGTTGTCATATTT 550
DB 1399 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1340
OY 551 ATCTTAAAAATGTGTAGCCAAATGATGTATGTTACATGTGAAAAAGTAGAACTGCTG 610
DB 1339 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1280
OY 611 AAAATGAAGATTAAATGCGCAGATTTGTTATTTCTAAAAAGCGTCAACAGCA 670
DB 1279 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1220
OY 671 CAGTAATAATATATCCAGAGTCAGATGATTAATATATATATATATATATATATATAT 730
DB 1219 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1160
OY 731 AAGCAAAACGCTTTATTAATGCGAAAAATTGAATATGCGGAAATGATTTCACAGAA 790
DB 1159 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1100
OY 791 CAGAGAAATGAAACGGAGAAATTAACAGTTAAAAATTTTGAAGT 833
DB 1099 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1057
```

RESULT 8

US-10-172-502-1
Sequence 1, Application US/10172502
Patent No. 6841154
GENERAL INFORMATION:
APPLICANT: FOSTER, Timothy et al.
TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.
FILE REFERENCE: P07263US01/BAS
CURRENT FILING DATE: 2002-06-17
PRIOR FILING DATE: 2001-06-15
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 6609
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-10-172-502-1

Query Match 3.4%; Score 64.6; DB 3; Length 6609;
Best Local Similarity 46.9%; Pred. No. 1.1e-05;

Matches 202; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

```
OY 545 ATATTATCTTAAATGTGTAGCCATGATGCTGTTTCAATGTGAAAAAAGTGGAA 604
DB 5549 ATTAAGAGAGCGACAGATGAAGAAACAATGCTGCAATAGTACAGTTGAAAAAGACTTAA 5608
OY 605 CTGCTGAATGAAGATTAAATGCGCAGAAATTGTTATTTCTTAAAGCGAGCTCAC 664
DB 5609 TTAAGCTTAAACAAATTTGCTGGGACGTAACTGATGTGGCATATTTATTTTCG 5668
```

```
OY 665 CAGGCACAGTAAATATATATCAAGAGTCAGAAAGATGATTAATATACATGCAACGAGTA 724
DB 5669 ATGATGGGAAAAACGAAATTCGTGAATCGAACCTGTATTAATTAATAAAGCAACTGCCG 5728
OY 725 AAGAACAGAACAAACGCTTATTAATTAATGCGAAAAAGTATGAAATTTGCGAAATGATTTCA 784
DB 5729 GAGAACATTTAACACATTAATTAACAGATTAAGAAACAGCAATTTGAAGCGAATGTTCAAG 5788
OY 785 CAGAACAGAGAAATGAGACGGAGAAATTAACAGTTAAAAATCTTGAGTTGCTGTATA 844
DB 5789 CAACAGTAAAGAAAGAAATAGTAATTTAGCAGATTAACAAAACATTTATGACACGTCTA 5848
OY 845 TTTTGAAGAAATGAAGTCCAAATTAATGCAAGATTAATGAAATCAACAAACAAAC 904
DB 5849 TTGGAACAAATTAATGATCAAGATCGTAGCAATGCAACAGTTGATTAACAGCAACTTAAATC 5908
OY 905 CATTTCATTAATGAAGCAACATCAACACCTGTTGAAAAAACAAGTCAAAATGATACCT 964
DB 5909 TACAAACATATACATGATTTTAGACGTACATCTTATTAATAAGCAGATGCTGAAAAACGA 5968
OY 965 CTAAAGTTGAT 975
DB 5969 TTAATGATGAT 5979
```

RESULT 9

US-09-902-540-1280
Sequence 1280, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: GOLDMAN, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT FILING DATE: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1280
LENGTH: 1039
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(1039)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

Query Match 3.0%; Score 57.2; DB 3; Length 1039;
Best Local Similarity 45.4%; Pred. No. 0.00036;

Matches 203; Conservative 0; Mismatches 244; Indels 0; Gaps 0;

```
OY 556 AAAATGTGTCGCCATATGATGTGTTTACATGCAAAAAAGTAGAACTGCTGAAT 615
DB 158 AAAAAAATTTAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 217
OY 616 GAAGATTAAATGCGCGCAATTTGTTATTTCTTAAAGCGAGCGCTCACAGCAGATGA 675
DB 218 TTTTAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 277
OY 676 AATATATCCAGAGTCAAGATGATTAATTAATCAATGACACAGATTAAGAAACAGCA 735
DB 278 AATTAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATA 337
OY 736 AAAAGCTTTATTAATGCGGAAAGTTATGAATTTGCGGAAATGATTTTCAAGAGAGAG 795
DB 338 TAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATA 397
OY 796 AATGAACGGAGAAATTAACAGTTAAAAATCTTGAGTTGTTGTTATTTTAAAGAA 855
```


Db 10614 ATGATGAGAAAAAGAAATTCGGAATGGAACCTGTATTAAACAGAAAGCCCTCGCTC 10555
Qy 725 AAGACAGACAAACCCCTTATTACTGGGAAAAGTTATGAAATGGCGAAATGATTCA 784
Db 10554 GAGAACAAATTCACAACTTATTCAACGATTAACAGCAATTTGAAGCGAATATTCAAG 10495
Qy 785 CAGAACAGAGATGGAACGGGAGAAATTAAACATTAAAACTTGAGCTGCTTCTGTAT 844
Db 10494 CAACGATGAGAAAGAAATAGTATATTGACAGATTACAAATTTATTGACACTGCTA 10435
Qy 845 TTTTGAAGAGTAAAGCTCCAAATATATGACAGAAATTAATTGAAAATCAACAAACAC 904
Db 10434 TTGACAAATTGATCAAGATCGTAGCAATGCAACAATTGATTAACAGCATCTTTAAATC 10375
Qy 905 CATTTCGAATTGAGCAACAAATCAACACCTGTTGAAAAAAGTCGCAAAAATGATACCT 964
Db 10374 TACAAACAATACATGATTGATGATCATCTATTAAAAAGCAGATGCTGAAAAACGA 10315
Qy 965 CTAAAGTTGAT 975
Db 10314 TTAAATGATGAT 10304

RESULT 12
US-08-781-986A-171/c
Sequence 171, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kuech
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 171:
SEQUENCE CHARACTERISTICS:
LENGTH: 11126 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-171

Query Match 3.0%; Score 56.6; DB 3; Length 11126;
Best Local Similarity 45.7%; Pred. No. 0.0011;
Matches 197; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

Qy 545 ATATTATCTTAAATGTGTGAGCCAAATGATGATGTTACATGTGAAAAAGTAGAGAA 604
Db 10734 ATAAAGAAAGCAAGATGAAGAAACAAATATTGCAATAGCAACAAGTTGAAAAAGATTAA 10675

Qy 605 CTGCTGAAATGAGATTTAAATGGCGAGAAATTGTTATTCTTAAAGCAAGCTCAC 664
Db 10674 TTAAACCTTAAACAAATTTGCTAGTGAAGTGAATCAATCAAGATGGCATATTATTTC 10615
Qy 665 CAGGCACAGTAAATATATCTCAAGAGTCAAAAGATGATTAATATACATGACAAACGATA 724
Db 10614 ATGATGAGAAAAAGAAATTCGGAATGGAACCTGTATTAAACAGAAAGCCCTCGCTC 10555
Qy 725 AAGACAGACAAACCCCTTATTACTGGGAAAAGTTATGAAATGGCGAAATGATTCA 784
Db 10554 GAGAACAAATTCACAACTTATTCAACGATTAACAGCAATTTGAAGCGAATATTCAAG 10495
Qy 785 CAGAACAGAGATGGAACGGGAGAAATTAAACATTAAAACTTGAGCTGCTTCTGTAT 844
Db 10494 CAACGATGAGAAAGAAATAGTATATTGACAGATTACAAATTTATTGACACTGCTA 10435
Qy 845 TTTTGAAGAGTAAAGCTCCAAATATATGACAGAAATTAATTGAAAATCAACAAACAC 904
Db 10434 TTGACAAATTGATCAAGATCGTAGCAATGCAACAATTGATTAACAGCATCTTTAAATC 10375
Qy 905 CATTTCGAATTGAGCAACAAATCAACACCTGTTGAAAAAAGTCGCAAAAATGATACCT 964
Db 10374 TACAAACAATACATGATTGATGATCATCTATTAAAAAGCAGATGCTGAAAAACGA 10315
Qy 965 CTAAAGTTGAT 975
Db 10314 TTAAATGATGAT 10304

RESULT 13
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bull et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco-
Patent No. 6503729
FILE OF INVENTION: jamaaschi1
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jamaaschi1
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature


```
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (10398)..(10398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (23180)..(23180)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234614)..(234614)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
```

```
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713552)..(713552)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

Query Match 3.0%; Score 56.4; DB 3; Length 1664976;
Best Local Similarity 50.0%; Pred.No. 0.0071;
Matches 141; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```
QY 833 TTGCTTCGATATTATTAGAGAGTAAAGCTCCAAATATGCAATTAATTGAAATC 892
Db 29123 TTGATTACAAACAAAGAGATGAGATGAAATTAATAAGAGTATTATATG 29064
QY 893 AAACAAAACACCATTTACATTTGAGCAACAAATCAACACCTGTGAAAAACAGTCA 952
```

Db 29063 AAAAAATTTAATAATCTATAGATAGAAAATATAGATAAAATGTAGAAAGATCCCTTG 29004
Qy 953 AAAAAATCTTAAAGTTGATTAACAACACCAAGCTTAGTGTAAAGATGTGGCA 1012
Db 29003 TAAATGCAAGCTTAATAAATTAATGAAAGAACGAAATTTTACAGCTTCAGAGAGGACG 28944
Qy 1013 TTGGCGAAAAATTAATATCAATTTCTGTAAATATTCATGCGGAGATTCAGACAAAG 1072
Db 28943 TTGCGATATTAACCACTATATATAAAATGAAACGAAAGATCACTAATATATGTGAAT 28884
Qy 1073 AAGCGACGCTAATAATAATACGTCAAATTCATTAGTGATA 1114
Db 28883 ATGGCTACTCAATATATCTTATACATTAAGTTCTGTAAAA 28842

RESULT 14
US-09-692-570-1/c
Sequence 1, Application US/09692570
Patent No. 6797466
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6797466
FILE REFERENCE: PB275C1
CURRENT APPLICATION NUMBER: US/09/692,570
CURRENT FILING DATE: 2003-01-14
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
PRIOR APPLICATION NUMBER: US 08/916,421
PRIOR FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g

FEATURE:
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:


```
Db 644 NNNNNNSCCTCTRMWMTMRWTKMGDGTVAKKYKWRDTCCTVYDVWADSVWVWYANMRCR 585
QY 704 TATATACATGACAAACGGATTAAGAACAAACGCTTATATCTGAGAAAAGTTATG 763
Db 584 DVTYTRMNTYCKSYASHSYWYGSNNMAMWYRYSARNWSMARWTR--NNMMWSGVRMR 528
QY 764 AAATTGGGGAATAATGATTTCACAGACGAGAGATGGAACGGGAAATTAACAGTTAAA 823
Db 527 WBGTMWRHRMNNNNNTDTRYYMMKKRABRTTYYDSMCNAKSWRGNRRAMKMMWAAAN 468
QY 824 ATCTGAGGTGGTTCGTATATTTAGAGAAAGTAAAGCTCCAAATAATGACAGATTAA 883
Db 467 DAGAMDHTWYWMGNNTMMWRBAKMMWMAWCRBAYCCNNNNNRACVWHGKMMRWTKYM 408
QY 884 TTGAAAATCAAAACAAAACACATTATCAATTGAGCAAAACAATCAACACCTGTGAAA 943
Db 407 WKACACNNNNBKAMMYRVAWWWYSRDTTNTDMMWMTSDMBMHWYTVDYTMRAAMNNNNNN 348
QY 944 AAACAGTCAAAAATGATACCTCTAAAGTGATAAACAACACCAAGCTTAGATGTAAG 1003
Db 347 WNBCKTTSWMMWMDHNTHTCTYGNNTWGSAYBMAAASMMWAGASNEVTYNNCMCRMTYMGK 288
QY 1004 ATGTGCAATTGGCGAAAAAATTAATATCAATTTCTGTAATATTCATTGGGGATTG 1063
Db 287 TMTNNNNNNKAWYRTKTVAMCENNRYYDTAVWTBKNYKCYAYBMYBYBMWKGHMBW 228
QY 1064 CAGACAAGAGCGACGCTAATAATACGTCAAATTCAATTAGTTAGTTAACAATGATG 1123
Db 227 WRRABHRSWNNMMWVCRCNKKYWSWHYHMYRBYKMAAVGCNNNNWKDRMAHHHWCATNNNM 168
QY 1124 CAGCCTTAACCTTTGATTAACGAGCTTCGAGAGATGCTATGCTATATGATGGGG 1183
Db 167 MWWWYAYMHHMKKGAAWNTNNKTABRDBHBAHVKTYYWRYDYWCAMCMNNAKAKAVRTA 108
QY 1184 ATACAGTATGCTCTGAAAATTATCAAGTACTGAACAAG 1225
Db 107 MGHMYYTDRYVSANNTGVMMWMMWRMCMWYSMMNNRWYRMG 66
```

Search completed: July 5, 2006, 22:11:16
Job time : 273 secs

THIS PAGE BLANK (USPTO)